GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 21, 2005, 12:46:16 ; Search time 5666.07 Seconds
(without alignments)
10732.529 Million cell updates/sec

Title: US-10-643-627-3
Perfect score: 1255
Sequence: 1 cgctccaggcctgggtgaca.....ttaagacctcctattgagtt 1255

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1255	100.0	1255	6	AR012638	AR012638 Sequence
	2	1255	100.0	1255	6	AR171258	AR171258 Sequence
	3	1255	100.0	1255	6	I42455	I42455 Sequence 3
	4	1255	100.0	1255	6	I87849	I87849 Sequence 3
	5	1247	99.4	18351	9	AF400075	AF400075 Homo sapi
c	6	1247	99.4	107278	9	AC114962	AC114962 Homo sapi
c	7	1247	99.4	184536	2	AC068682	AC068682 Homo sapi
	8	1243.8	99.1	52368	9	AC010621	AC010621 Homo sapi
	9	1211	96.5	1289	9	HSPAR2B	Z49994 H.sapiens p
	10	1117	89.0	1451	6	CQ726252	CQ726252 Sequence
	11	1117	89.0	2813	9	BC012453	BC012453 Homo sapi
	12	1115.4	88.9	2876	9	BC018130	BC018130 Homo sapi
	13	1114	88.8	1275	9	AY336105	AY336105 Homo sapi
	14	1113.8	88.7	1451	6	CQ870621	CQ870621 Sequence
	15	1113.8	88.7	1451	6	CQ876755	CQ876755 Sequence
	16	1113.8	88.7	1451	6	AX549014	AX549014 Sequence
	17	1113.8	88.7	1451	9	HSU34038	U34038 Human prote
	18	1112	88.6	1194	9	BT009856	BT009856 Homo sapi
	19	1105.8	88.1	1414	6	AR012640	AR012640 Sequence
	20	1105.8	88.1	1414	6	AR171260	AR171260 Sequence
	21	1105.8	88.1	1414	6	I87851	I87851 Sequence 62
	22	1105.4	88.1	1124	9	HSU36753	U36753 Human prote
	23	1066.4	85.0	1068	6	CQ760412	CQ760412 Sequence
	24	1024.4	81.6	1026	6	CQ760413	CQ760413 Sequence
	25	961.4	76.6	963	6	CQ760414	CQ760414 Sequence
	26	845.8	67.4	1477	10	MMPROACR2	Z35158 M.musculus
c	27	845.8	67.4	199227	2	AC110876	AC110876 Mus muscu
	28	844.2	67.3	1475	6	AR012637	AR012637 Sequence
	29	844.2	67.3	1475	6	AR171257	AR171257 Sequence
	30	844.2	67.3	1475	6	I42454	I42454 Sequence 1
	31	844.2	67.3	1475	6	I87848	I87848 Sequence 1
	32	823.8	65.6	2772	10	BC025432	BC025432 Mus muscu
	33	817.4	65.1	2713	6	AX785009	AX785009 Sequence
	34	817.4	65.1	2713	10	MMPAR2MR	Z48043 M.musculus
	35	817.4	65.1	2732	6	AR012639	AR012639 Sequence
	36	817.4	65.1	2732	6	AR171259	AR171259 Sequence
	37	817.4	65.1	2732	6	I87850	I87850 Sequence 60
c	38	805.4	64.2	175653	2	AC128225	AC128225 Rattus no
	39	805.4	64.2	182064	2	AC130634	AC130634 Rattus no
	40	790.2	63.0	1428	10	RNU61373	U61373 Rattus norv
	41	673.2	53.6	846	11	BV166225	BV166225 F2RL1_620
	42	575.6	45.9	91325	2	AC139936	AC139936 Mus muscu
c	43	570.4	45.5	91325	2	AC139936	AC139936 Mus muscu
	44	341.8	27.2	180416	5	BX323555	BX323555 Zebrafish
	45	341.8	27.2	194749	2	BX842690	BX842690 Danio rer

ALIGNMENTS

RESULT 1
AR012638
LOCUS AR012638 1255 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 3 from patent US 5763575.
ACCESSION AR012638
VERSION AR012638.1 GI:3970628
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1255)
AUTHORS Sundelin,J. and Scarborough,R.M.
TITLE Agonist and antagonist peptides of the C140 receptor
JOURNAL Patent: US 5763575-A 3 09-JUN-1998;
FEATURES Location/Qualifiers
source 1. .1255
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 1255; DB 6; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGCTCCAGGCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCTCCAGGCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAA 60

Qy         61 TGTACTTTCATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTACTTTCATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCT 120

Qy        121 TGTCTTCCTTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTCTTCCTTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAA 180

Qy        181 GGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGT 240

Qy        241 GGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 GGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGT 300

Qy        301 CTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCT 360

Qy        361 TTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGA 420

Qy        421 CCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTG 480

Qy        481 GATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTG 540

Qy        541 TTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCAT 600

Qy        601 GGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCT 660

Qy        661 GATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCT 720

Qy        721 GAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAA 780

Qy        781 TTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGC 840

Qy        841 CTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAA 900

Qy        901 AAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCAC 960

Qy        961 TCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCA 1020

Qy       1021 TGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCC 1080

Qy       1081 CTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTG 1140

Qy       1141 CCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAG 1200

Qy       1201 GAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255


RESULT 2
AR171258
LOCUS       AR171258               1255 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 3 from patent US 6297026.
ACCESSION   AR171258
VERSION     AR171258.1  GI:17910208
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1255)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Nucleic acids encoding the C140 receptor
  JOURNAL   Patent: US 6297026-A 3 02-OCT-2001;
FEATURES             Location/Qualifiers
     source          1. .1255
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 1255;  DB 6;  Length 1255;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGCTCCAGGCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCTCCAGGCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAA 60

Qy         61 TGTACTTTCATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTACTTTCATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCT 120

Qy        121 TGTCTTCCTTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTCTTCCTTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAA 180

Qy        181 GGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGT 240

Qy        241 GGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGT 300

Qy        301 CTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCT 360

Qy        361 TTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGA 420

Qy        421 CCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTG 480

Qy        481 GATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTG 540

Qy        541 TTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCAT 600

Qy        601 GGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCT 660

Qy        661 GATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCT 720

Qy        721 GAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAA 780

Qy        781 TTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGC 840
```

```
Qy        841 CTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAA 900

Qy        901 AAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCAC 960

Qy        961 TCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCA 1020

Qy       1021 TGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCC 1080

Qy       1081 CTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTG 1140

Qy       1141 CCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAG 1200

Qy       1201 GAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255


RESULT 3
I42455
LOCUS       I42455                  1255 bp    DNA     linear   PAT 07-OCT-1997
DEFINITION  Sequence 3 from patent US 5629174.
ACCESSION   I42455
VERSION     I42455.1  GI:2467950
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1255)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Recombinant C140 receptor
  JOURNAL   Patent: US 5629174-A 3 13-MAY-1997;
FEATURES             Location/Qualifiers
     source          1. .1255
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 1255;  DB 6;  Length 1255;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGCTCCAGGCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCTCCAGGCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAA 60

Qy         61 TGTACTTTCATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTACTTTCATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCT 120

Qy        121 TGTCTTCCTTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTCTTCCTTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAA 180

Qy        181 GGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGT 240

Qy        241 GGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGT 300

Qy        301 CTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCT 360

Qy        361 TTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGA 420

Qy        421 CCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTG 480

Qy        481 GATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTG 540

Qy        541 TTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCAT 600

Qy        601 GGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCT 660

Qy        661 GATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCT 720

Qy        721 GAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAA 780

Qy        781 TTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGC 840

Qy        841 CTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAA 900

Qy        901 AAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCAC 960

Qy        961 TCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCA 1020

Qy       1021 TGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCC 1080

Qy       1081 CTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTG 1140

Qy       1141 CCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAG 1200

Qy       1201 GAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255


RESULT 4
I87849
LOCUS       I87849                  1255 bp    DNA     linear   PAT 10-AUG-1998
DEFINITION  Sequence 3 from patent US 5716789.
ACCESSION   I87849
VERSION     I87849.1  GI:3407789
KEYWORDS    .
```

```
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1255)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Method to determine ligands, agonist and antagonist of C140
            receptor
  JOURNAL   Patent: US 5716789-A 3 10-FEB-1998;
FEATURES             Location/Qualifiers
     source          1. .1255
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

Query Match 100.0%; Score 1255; DB 6; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGCTCCAGGCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCTCCAGGCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAA 60

Qy         61 TGTACTTTCATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTACTTTCATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCT 120

Qy        121 TGTCTTCCTTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTCTTCCTTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAA 180

Qy        181 GGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGT 240

Qy        241 GGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGT 300

Qy        301 CTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCT 360

Qy        361 TTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGA 420

Qy        421 CCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTG 480

Qy        481 GATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTG 540

Qy        541 TTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCAT 600

Qy        601 GGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCT 660

Qy        661 GATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCT 720

Qy        721 GAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAA 780

Qy        781 TTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGC 840

Qy        841 CTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAA 900

Qy        901 AAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCAC 960

Qy        961 TCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCA 1020

Qy       1021 TGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCC 1080

Qy       1081 CTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTG 1140

Qy       1141 CCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAG 1200

Qy       1201 GAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
```

```
RESULT 5
AF400075
LOCUS       AF400075               18351 bp    DNA     linear   PRI 18-JAN-2002
DEFINITION  Homo sapiens coagulation factor II (thrombin) receptor-like 1
            (F2RL1) gene, complete cds.
ACCESSION   AF400075
VERSION     AF400075.1  GI:15021772
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 18351)
  AUTHORS   Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
            Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2001) Molecular Biotechnology, University of
            Washington, 1705 NE Pacific, Seattle, WA 98195, USA
COMMENT     To cite this work please use:  SeattleSNPs. NHLBI Program for
            Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.mbt.washington.edu)
            This sequence consists of 2 contigs.  The gap between the contigs
            is represented as a run of N.
                1    10774: contig of 10774 bp in length
            10775    10874: gap of unknown length
            10875    18351: contig of 7477 bp in length.
FEATURES             Location/Qualifiers
     source          1. .18351
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     repeat_region   343. .492
                     /rpt_family="MIR"
                     /rpt_type=dispersed
     variation       467
                     /frequency="0.09"
                     /replace="a"
     variation       531
                     /frequency="0.01"
                     /replace="g"
     variation       678
                     /frequency="0.01"
                     /replace="c"
```

```
repeat_region   728. .780
                /rpt_family="L2"
                /rpt_type=dispersed
variation       829
                /frequency="0.03"
                /replace="a"
repeat_region   913. .1028
                /rpt_family="FLAM_C"
                /rpt_type=dispersed
variation       925
                /frequency="0.01"
                /replace="t"
repeat_region   1059. .1214
                /rpt_family="L3"
                /rpt_type=dispersed
repeat_region   1292. .1578
                /rpt_family="AluSg"
                /rpt_type=dispersed
variation       1538
                /frequency="0.01"
                /replace="g"
variation       1622
                /frequency="0.02"
                /replace="t"
variation       1623
                /frequency="0.01"
                /replace="g"
repeat_region   1795. .1926
                /rpt_family="MIR"
                /rpt_type=dispersed
variation       2122
                /frequency="0.01"
                /replace="g"
variation       2254
                /frequency="0.01"
                /replace="c"
variation       2457
                /frequency="0.10"
                /replace="g"
gene            2462. .17176
                /gene="F2RL1"
mRNA            join(2462. .2687,15958. .17176)
                /gene="F2RL1"
                /product="coagulation factor II (thrombin) receptor-like
                1"
variation       2561
                /gene="F2RL1"
                /frequency="0.28"
                /replace="c"
CDS             join(2606. .2687,15958. .17069)
                /gene="F2RL1"
                /codon_start=1
                /product="coagulation factor II (thrombin) receptor-like
                1"
                /protein_id="AAK77914.1"
                /db_xref="GI:15021773"
                /translation="MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG
                TSHVTGKGVTVETVFSVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLF
                RTKKKHPAVIYMANLALADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMY
                CSILFMTCLSVQRYWVIVNPMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFI
                PALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDE
                NSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLST
                LNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTT
                VKTSY"
variation       2667
                /gene="F2RL1"
                /frequency="0.02"
                /replace="t"
variation       2770
                /gene="F2RL1"
                /frequency="0.09"
                /replace="a"
variation       3069
                /gene="F2RL1"
                /frequency="0.03"
                /replace="a"
variation       3119
                /gene="F2RL1"
                /frequency="0.32"
                /replace="t"
variation       3151
                /gene="F2RL1"
                /frequency="0.01"
                /replace="g"
variation       3263
                /gene="F2RL1"
                /frequency="0.33"
                /replace="g"
variation       3306
                /gene="F2RL1"
                /frequency="0.02"
                /replace="t"
variation       3330
                /gene="F2RL1"
                /frequency="0.01"
                /replace="c"
variation       3331
                /gene="F2RL1"
                /frequency="0.07"
                /replace="t"
variation       3353
                /gene="F2RL1"
                /frequency="0.01"
                /replace="a"
variation       3417
                /gene="F2RL1"
                /frequency="0.01"
                /replace="c"
variation       4160
                /gene="F2RL1"
                /frequency="0.01"
                /replace="c"
variation       4456
                /gene="F2RL1"
                /frequency="0.01"
                /replace=""
variation       4500
                /gene="F2RL1"
                /frequency="0.02"
                /replace="t"
variation       4663
                /gene="F2RL1"
                /frequency="0.46"
                /replace="t"
repeat_region   4710. .4817
                /rpt_family="L2"
                /rpt_type=dispersed
variation       4715
                /gene="F2RL1"
                /frequency="0.15"
                /replace="g"
variation       4972
                /gene="F2RL1"
                /frequency="0.02"
                /replace="c"
repeat_region   4993. .5299
                /rpt_family="AluSx"
                /rpt_type=dispersed
variation       5065
                /gene="F2RL1"
                /frequency="0.01"
                /replace="g"
variation       5159
                /gene="F2RL1"
                /frequency="0.01"
                /replace=""
```

```
     variation       5295
                     /gene="F2RL1"
                     /frequency="0.01"
                     /replace="t"
     variation       5296
                     /gene="F2RL1"
                     /frequency="0.01"
                     /replace="a"
     variation       5366
                     /gene="F2RL1"
                     /frequency="0.17"
                     /replace="t"
     repeat_region   5643. .5940

  Query Match             99.4%;  Score 1247;  DB 9;  Length 18351;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 GCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAATGTACTTT 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15826 GCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAATGTACTTT 15885

Qy         69 CATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCTTGTCTTCC 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15886 CATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCTTGTCTTCC 15945

Qy        129 TTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATG 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15946 TTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATG 16005

Qy        189 GCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGT 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16006 GCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGT 16065

Qy        249 TTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAA 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16066 TTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAA 16125

Qy        309 TTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAA 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16126 TTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAA 16185

Qy        369 CTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCT 428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16186 CTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCT 16245

Qy        429 CTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATG 488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16246 CTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATG 16305

Qy        489 GGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTC 548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16306 GGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTC 16365

Qy        549 TCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACT 608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16366 TCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACT 16425

Qy        609 CCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGC 668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16426 CCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGC 16485

Qy        669 TGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCA 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16486 TGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCA 16545

Qy        729 CGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCC 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16546 CGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCC 16605

Qy        789 TCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGC 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16606 TCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGC 16665

Qy        849 TGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGA 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16666 TGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGA 16725

Qy        909 GGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTA 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16726 GGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTA 16785

Qy        969 ACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATG 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16786 ACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATG 16845

Qy       1029 CCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCT 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16846 CCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCT 16905

Qy       1089 ATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTG 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16906 ATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTG 16965

Qy       1149 TCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCA 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16966 TCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCA 17025

Qy       1209 GCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db      17026 GCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 17072


RESULT 6
AC114962/c
LOCUS       AC114962               107278 bp    DNA     linear   PRI 05-NOV-2002
DEFINITION  Homo sapiens chromosome 5 clone RP11-206N2, complete sequence.
ACCESSION   AC114962
VERSION     AC114962.2  GI:24580362
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107278)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 107278)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 107278)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Nov 5, 2002 this sequence version replaced gi:19424423.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.7% of Sequence;
            Estimated Total Number of Errors is 0.2.
            NOTE:  This insert is not the entire sequence of the clone (entire
            sequence is 201.6kb). It is clipped at the overlaps with AC027342
            and AC010621. The number of bases overlapped with AC027342 is 5439
            and with AC010621 is 28528.
FEATURES             Location/Qualifiers
     source          1. .107278
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-206N2"
```

ORIGIN

Query Match 99.4%; Score 1247; DB 9; Length 107278;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          9 GCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAATGTACTTT 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87712 GCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAATGTACTTT 87653

Qy         69 CATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCTTGTCTTCC 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87652 CATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCTTGTCTTCC 87593

Qy        129 TTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATG 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87592 TTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATG 87533

Qy        189 GCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGT 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87532 GCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGT 87473

Qy        249 TTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAA 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87472 TTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAA 87413

Qy        309 TTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAA 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87412 TTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAA 87353

Qy        369 CTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCT 428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87352 CTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCT 87293

Qy        429 CTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATG 488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87292 CTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATG 87233

Qy        489 GGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTC 548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87232 GGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTC 87173

Qy        549 TCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACT 608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87172 TCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACT 87113

Qy        609 CCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGC 668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87112 CCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGC 87053

Qy        669 TGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCA 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87052 TGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCA 86993

Qy        729 CGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCC 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86992 CGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCC 86933

Qy        789 TCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGC 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86932 TCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGC 86873

Qy        849 TGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGA 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86872 TGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGA 86813

Qy        909 GGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTA 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86812 GGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTA 86753

Qy        969 ACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATG 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86752 ACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATG 86693

Qy       1029 CCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCT 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86692 CCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCT 86633

Qy       1089 ATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTG 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86632 ATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTG 86573

Qy       1149 TCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCA 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86572 TCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCA 86513

Qy       1209 GCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
              |||||||||||||||||||||||||||||||||||||||||||||||
Db      86512 GCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 86466
```

RESULT 7
AC068682/c
LOCUS AC068682 184536 bp DNA linear HTG 23-SEP-2000
DEFINITION Homo sapiens clone RP11-206N2, WORKING DRAFT SEQUENCE, 37 unordered pieces.
ACCESSION AC068682
VERSION AC068682.3 GI:10280868
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 184536)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-206N2
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 184536)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Sep 23, 2000 this sequence version replaced gi:8247861.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
---------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
---------- Project Information
Center project name: L5802
Center clone name: 206_N_2
---------- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads

```
    Chemistry: Dye-terminator Big Dye; 100% of reads
    Assembly program: Phrap; version 0.960731
    Consensus quality: 167659 bases at least Q40
    Consensus quality: 175736 bases at least Q30
    Consensus quality: 178914 bases at least Q20
    Insert size: 154000; agarose-fp
    Insert size: 180936; sum-of-contigs
    Quality coverage: 5.0 in Q20 bases; agarose-fp
    Quality coverage: 4.3 in Q20 bases; sum-of-contigs
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 37 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1154: contig of 1154 bp in length
*     1155     1254: gap of 100 bp
*     1255     2766: contig of 1512 bp in length
*     2767     2866: gap of 100 bp
*     2867     4073: contig of 1207 bp in length
*     4074     4173: gap of 100 bp
*     4174     5191: contig of 1018 bp in length
*     5192     5291: gap of 100 bp
*     5292     7387: contig of 2096 bp in length
*     7388     7487: gap of 100 bp
*     7488     9087: contig of 1600 bp in length
*     9088     9187: gap of 100 bp
*     9188    10361: contig of 1174 bp in length
*    10362    10461: gap of 100 bp
*    10462    12065: contig of 1604 bp in length
*    12066    12165: gap of 100 bp
*    12166    13498: contig of 1333 bp in length
*    13499    13598: gap of 100 bp
*    13599    15085: contig of 1487 bp in length
*    15086    15185: gap of 100 bp
*    15186    17249: contig of 2064 bp in length
*    17250    17349: gap of 100 bp
*    17350    19243: contig of 1894 bp in length
*    19244    19343: gap of 100 bp
*    19344    21127: contig of 1784 bp in length
*    21128    21227: gap of 100 bp
*    21228    23345: contig of 2118 bp in length
*    23346    23445: gap of 100 bp
*    23446    25965: contig of 2520 bp in length
*    25966    26065: gap of 100 bp
*    26066    28195: contig of 2130 bp in length
*    28196    28295: gap of 100 bp
*    28296    32216: contig of 3921 bp in length
*    32217    32316: gap of 100 bp
*    32317    36954: contig of 4638 bp in length
*    36955    37054: gap of 100 bp
*    37055    40348: contig of 3294 bp in length
*    40349    40448: gap of 100 bp
*    40449    43556: contig of 3108 bp in length
*    43557    43656: gap of 100 bp
*    43657    47926: contig of 4270 bp in length
*    47927    48026: gap of 100 bp
*    48027    54079: contig of 6053 bp in length
*    54080    54179: gap of 100 bp
*    54180    58962: contig of 4783 bp in length
*    58963    59062: gap of 100 bp
*    59063    63739: contig of 4677 bp in length
*    63740    63839: gap of 100 bp
*    63840    71551: contig of 7712 bp in length
*    71552    71651: gap of 100 bp
*    71652    78218: contig of 6567 bp in length
*    78219    78318: gap of 100 bp
*    78319    84982: contig of 6664 bp in length
*    84983    85082: gap of 100 bp
*    85083    91930: contig of 6848 bp in length
*    91931    92030: gap of 100 bp
*    92031    98922: contig of 6892 bp in length
*    98923    99022: gap of 100 bp
*    99023   106496: contig of 7474 bp in length
*   106497   106596: gap of 100 bp
*   106597   114950: contig of 8354 bp in length
*   114951   115050: gap of 100 bp
*   115051   124353: contig of 9303 bp in length
*   124354   124453: gap of 100 bp
*   124454   132639: contig of 8186 bp in length
*   132640   132739: gap of 100 bp
*   132740   141572: contig of 8833 bp in length
*   141573   141672: gap of 100 bp
*   141673   151302: contig of 9630 bp in length
*   151303   151402: gap of 100 bp
*   151403   161767: contig of 10365 bp in length
*   161768   161867: gap of 100 bp
*   161868   184536: contig of 22669 bp in length.
FEATURES             Location/Qualifiers
     source          1. .184536
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-206N2"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1154
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    1255. .2766
                     /note="assembly_fragment"
     misc_feature    2867. .4073
                     /note="assembly_fragment"
     misc_feature    4174. .5191
                     /note="assembly_fragment"
     misc_feature    5292. .7387
                     /note="assembly_fragment"
     misc_feature    7488. .9087
                     /note="assembly_fragment"
     misc_feature    9188. .10361
                     /note="assembly_fragment"
     misc_feature    10462. .12065
                     /note="assembly_fragment"
     misc_feature    12166. .13498
                     /note="assembly_fragment"
     misc_feature    13599. .15085
                     /note="assembly_fragment"
     misc_feature    15186. .17249
                     /note="assembly_fragment"
     misc_feature    17350. .19243
                     /note="assembly_fragment"
     misc_feature    19344. .21127
                     /note="assembly_fragment"
     misc_feature    21228. .23345
                     /note="assembly_fragment"
     misc_feature    23446. .25965
                     /note="assembly_fragment"
     misc_feature    26066. .28195
                     /note="assembly_fragment"
     misc_feature    28296. .32216
                     /note="assembly_fragment"
     misc_feature    32317. .36954
                     /note="assembly_fragment"
     misc_feature    37055. .40348
                     /note="assembly_fragment"
     misc_feature    40449. .43556
                     /note="assembly_fragment"
     misc_feature    43657. .47926
                     /note="assembly_fragment"
     misc_feature    48027. .54079
                     /note="assembly_fragment"
     misc_feature    54180. .58962
                     /note="assembly_fragment"
```

```
     misc_feature    59063. .63739

  Query Match             99.4%;  Score 1247;  DB 2;  Length 184536;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 GCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAATGTACTTT 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66894 GCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAATGTACTTT 66835

Qy         69 CATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCTTGTCTTCC 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66834 CATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCTTGTCTTCC 66775

Qy        129 TTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATG 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66774 TTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATG 66715

Qy        189 GCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGT 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66714 GCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGT 66655

Qy        249 TTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAA 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66654 TTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAA 66595

Qy        309 TTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAA 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66594 TTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAA 66535

Qy        369 CTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCT 428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66534 CTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCT 66475

Qy        429 CTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATG 488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66474 CTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATG 66415

Qy        489 GGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTC 548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66414 GGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTC 66355

Qy        549 TCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACT 608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66354 TCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACT 66295

Qy        609 CCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGC 668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66294 CCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGC 66235

Qy        669 TGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCA 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66234 TGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCA 66175

Qy        729 CGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCC 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66174 CGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCC 66115

Qy        789 TCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGC 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66114 TCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGC 66055

Qy        849 TGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGA 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66054 TGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGA 65995

Qy        909 GGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTA 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65994 GGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTA 65935

Qy        969 ACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATG 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65934 ACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATG 65875

Qy       1029 CCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCT 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65874 CCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCT 65815

Qy       1089 ATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTG 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65814 ATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTG 65755

Qy       1149 TCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCA 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65754 TCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCA 65695

Qy       1209 GCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
              |||||||||||||||||||||||||||||||||||||||||||||||
Db      65694 GCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 65648


RESULT 8
AC010621
LOCUS       AC010621               52368 bp    DNA     linear   PRI 26-JAN-2002
DEFINITION  Homo sapiens chromosome 5 clone CTD-2034J20, complete sequence.
ACCESSION   AC010621
VERSION     AC010621.7  GI:18376859
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 52368)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 52368)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-SEP-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 52368)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jan 26, 2002 this sequence version replaced gi:13699549.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.4% of Sequence;
            Estimated Total Number of Errors is 0.2.
            NOTE:  This insert is not the entire sequence of the clone (entire
            sequence is 142.2kb). It is clipped at the overlap with AC025188.
            The number of bases overlapped is 12615.
FEATURES             Location/Qualifiers
     source          1. .52368
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2034J20"
ORIGIN

  Query Match             99.1%;  Score 1243.8;  DB 9;  Length 52368;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1245;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          9 GCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAATGTACTTT 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43406 GCCTGGGTGACAGCGAGACCCTGTCTCATAAATTAAAAAATGAATAAATGAATGTACTTT 43465

Qy         69 CATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCTTGTCTTCC 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43466 CATTTGAACAAACCAGTGTTACTGCTGAAACATTTATTTCTGTAATGACCCTTGTCTTCC 43525
```

```
Qy       129 TTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATG 188
            ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     43526 TTTCTTGTACAGGAACCAGTAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATG 43585

Qy       189 GCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGT 248
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     43586 GCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGT 43645

Qy       249 TTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAA 308
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     43646 TTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAA 43705

Qy       309 TTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAA 368
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     43706 TTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAA 43765

Qy       369 CTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCT 428
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     43766 CTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCT 43825

Qy       429 CTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATG 488
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     43826 CTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATG 43885

Qy       489 GGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTC 548
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     43886 GGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTC 43945

Qy       549 TCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACT 608
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     43946 TCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACT 44005

Qy       609 CCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGC 668
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44006 CCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGC 44065

Qy       669 TGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCA 728
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44066 TGGTCACCATTCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCA 44125

Qy       729 CGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCC 788
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44126 CGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCC 44185

Qy       789 TCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGC 848
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44186 TCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGC 44245

Qy       849 TGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGA 908
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44246 TGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGA 44305

Qy       909 GGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTA 968
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44306 GGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTA 44365

Qy       969 ACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATG 1028
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44366 ACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATG 44425

Qy      1029 CCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCT 1088
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44426 CCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCT 44485

Qy      1089 ATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTG 1148
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44486 ATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTG 44545

Qy      1149 TCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCA 1208
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44546 TCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCA 44605
```

```
Qy      1209 GCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
            |||||||||||||||||||||||||||||||||||||||||||||||
Db     44606 GCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 44652
```

RESULT 9
HSPAR2B
LOCUS HSPAR2B 1289 bp DNA linear PRI 09-SEP-2004
DEFINITION H.sapiens partial gene for proteinase-activated receptor 2 (1289 BP).
ACCESSION Z49994
VERSION Z49994.1 GI:1008086
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1289)
AUTHORS Nystedt,S., Emilsson,K., Larsson,A.K., Strombeck,B. and Sundelin,J.
TITLE Molecular cloning and functional expression of the gene encoding the human proteinase-activated receptor 2
JOURNAL Eur. J. Biochem. 232 (1), 84-89 (1995)
MEDLINE 96048032
PUBMED 7556175
REFERENCE 2 (bases 1 to 1289)
AUTHORS Nystedt,S.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-1995) Sverker Nystedt, Division of Neurobiology, The Wallenberg, Laboratory, Lund University, Soelvegatan 33A, Lund, S-223 62, Sweden
FEATURES Location/Qualifiers
source 1. .1289
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/map="q13"
intron <1. .96
exon 97. .>1289
/product="proteinase-activated receptor 2"
/usedin=Z49993:par2_CDS
3'UTR 1209. .>1289
ORIGIN

Query Match 96.5%; Score 1211; DB 9; Length 1289;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        45 AAAATGAATAAATGAATGTACTTTCATTTGAACAAACCAGTGTTACTGCTGAAACATTTA 104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 AAAATGAATAAATGAATGTACTTTCATTTGAACAAACCAGTGTTACTGCTGAAACATTTA 60

Qy       105 TTTCTGTAATGACCCTTGTCTTCCTTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAA 164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TTTCTGTAATGACCCTTGTCTTCCTTTCTTGTACAGGAACCAATAGATCCTCTAAAGGAA 120

Qy       165 GAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTG 224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTG 180

Qy       225 AAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACTG 284
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 AAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACTG 240

Qy       285 TCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGG 344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGG 300

Qy       345 CCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCA 404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCA 360

Qy       405 ATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACA 464
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACA 420

Qy        465 TACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCT 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCT 480

Qy        525 ATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGG 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGG 540

Qy        585 TCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCC 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCC 600

Qy        645 TGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCA 704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCA 660

Qy        705 TCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGG 764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGG 720

Qy        765 TGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCT 824
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCT 780

Qy        825 TCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATG 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATG 840

Qy        885 AAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGT 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGT 900

Qy        945 ACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGA 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGA 960

Qy       1005 GCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTA 1064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTA 1020

Qy       1065 ACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAA 1124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAA 1080

Qy       1125 AGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCT 1184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCT 1140

Qy       1185 CAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCT 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCT 1200

Qy       1245 CCTATTGAGTT 1255
              |||||||||||
Db       1201 CCTATTGAGTT 1211
```

```
RESULT 10
CQ726252
LOCUS       CQ726252                1451 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 12186 from Patent WO02068579.
ACCESSION   CQ726252
VERSION     CQ726252.1  GI:42288766
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 12186 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1..1451
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             89.0%;  Score 1117;  DB 6;  Length 1451;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-306;
  Matches 1117;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        139 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 288

Qy        199 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 348

Qy        259 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 408

Qy        319 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 468

Qy        379 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 528

Qy        439 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 588

Qy        499 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 648

Qy        559 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 708

Qy        619 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 768

Qy        679 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 828

Qy        739 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 888

Qy        799 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 948

Qy        859 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 1008

Qy        919 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 1068

Qy        979 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1038
```

```
Db      1069 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1128

Qy      1039 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1098

Db      1129 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1188

Qy      1099 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1158

Db      1189 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1248

Qy      1159 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1218

Db      1249 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1308

Qy      1219 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255

Db      1309 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1345
```

RESULT 11
BC012453

```
LOCUS       BC012453                2813 bp    mRNA    linear   PRI 30-JUN-2004
DEFINITION  Homo sapiens coagulation factor II (thrombin) receptor-like 1, mRNA
            (cDNA clone MGC:9549 IMAGE:3857382), complete cds.
ACCESSION   BC012453
VERSION     BC012453.1  GI:15214649
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2813)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2813)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 21 Row: b Column: 20
            This clone was selected for full length sequencing because it
            passed the following selection criteria: GenomeScan gene
            prediction, Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1. .2813
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:9549 IMAGE:3857382"
                     /tissue_type="Ovary, adenocarcinoma"
                     /clone_lib="NIH_MGC_66"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .2813
                     /gene="F2RL1"
                     /note="synonyms: GPR11, PAR2"
                     /db_xref="LocusID:2150"
                     /db_xref="MIM:600933"
     CDS             93. .1286
                     /gene="F2RL1"
                     /codon_start=1
                     /product="coagulation factor II (thrombin) receptor-like
                     1, precursor"
                     /protein_id="AAH12453.1"
                     /db_xref="GI:15214650"
                     /db_xref="LocusID:2150"
                     /db_xref="MIM:600933"
                     /translation="MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG
                     TSHVTGKGVTVETVFSVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLF
                     RTKKKHPAVIYMANLALADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMY
                     CSILFMTCLSVQRYWVIVNPMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFI
                     PALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDE
                     NSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLST
                     LNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTT
                     VKTSY"
ORIGIN

  Query Match             89.0%;  Score 1117;  DB 9;  Length 2813;
  Best Local Similarity   100.0%;  Pred. No. 8e-306;
  Matches 1117;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       139 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 198

Db       173 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 232

Qy       199 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 258

Db       233 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 292

Qy       259 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 318

Db       293 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 352

Qy       319 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 378

Db       353 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 412

Qy       379 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 438

Db       413 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 472

Qy       439 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 498

Db       473 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 532

Qy       499 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 558
```

```
Db        533 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 592

Qy        559 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 652

Qy        619 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 712

Qy        679 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 772

Qy        739 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 832

Qy        799 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 892

Qy        859 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 952

Qy        919 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 1012

Qy        979 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1072

Qy       1039 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1132

Qy       1099 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1192

Qy       1159 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1252

Qy       1219 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
              |||||||||||||||||||||||||||||||||||||
Db       1253 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1289


RESULT 12
BC018130
LOCUS       BC018130               2876 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens coagulation factor II (thrombin) receptor-like 1, mRNA
            (cDNA clone MGC:9298 IMAGE:3895653), complete cds.
ACCESSION   BC018130
VERSION     BC018130.1  GI:17390291
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2876)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2876)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-DEC-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 14 Row: h Column: 12
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 8051581.
FEATURES             Location/Qualifiers
     source          1. .2876
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:9298 IMAGE:3895653"
                     /tissue_type="Pancreas, epithelioid carcinoma"
                     /clone_lib="NIH_MGC_70"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .2876
                     /gene="F2RL1"
                     /note="synonyms: GPR11, PAR2"
                     /db_xref="LocusID:2150"
                     /db_xref="MIM:600933"
     CDS             155. .1348
                     /gene="F2RL1"
                     /codon_start=1
                     /product="coagulation factor II (thrombin) receptor-like
                     1, precursor"
                     /protein_id="AAH18130.1"
                     /db_xref="GI:17390292"
                     /db_xref="LocusID:2150"
                     /db_xref="MIM:600933"
                     /translation="MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG
                     TSHVTGKGVTVETVFSVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLF
                     RTKKKHPAVIYMANLALADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMY
                     CSILFMTCLSVQRYWVIVNPMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFI
                     PALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDE
                     NSEKKRKRAIKLIVSVLAMYLICFTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLST
                     LNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTT
                     VKTSY"
ORIGIN
```

```
  Query Match            88.9%;  Score 1115.4;  DB 9;  Length 2876;
  Best Local Similarity   99.9%;  Pred. No. 2.3e-305;
  Matches 1116;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        139 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 294

Qy        199 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 354

Qy        259 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 414

Qy        319 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 474

Qy        379 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 534

Qy        439 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 594

Qy        499 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 654

Qy        559 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 714

Qy        619 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 774

Qy        679 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 834

Qy        739 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 894

Qy        799 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 954

Qy        859 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 1014

Qy        919 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 978
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 ACTCATTGTCTCTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 1074

Qy        979 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1134

Qy       1039 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1194

Qy       1099 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1254

Qy       1159 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1314

Qy       1219 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
              |||||||||||||||||||||||||||||||||||||
Db       1315 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1351
```

```
RESULT 13
AY336105
LOCUS       AY336105                1275 bp    mRNA    linear   PRI 28-JUL-2003
DEFINITION  Homo sapiens protease-activated receptor 2 (F2RL1) mRNA, complete
            cds.
ACCESSION   AY336105
VERSION     AY336105.1  GI:33149991
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1275)
  AUTHORS   Luo,W., Sedehizade,F., Hanck,T. and Reiser,G.
  TITLE     Human protease-activated receptor 2 (PAR-2)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1275)
  AUTHORS   Luo,W., Sedehizade,F., Hanck,T. and Reiser,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2003) Medical Faculty, Institute of
            Neurobiochemistry, Leipziger Str. 44, Magdeburg, Saxony-Anhalt
            39120, Germany
FEATURES             Location/Qualifiers
     source          1. .1275
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="5q13"
                     /cell_line="HEK293; A549"
     gene            1. .1275
                     /gene="F2RL1"
                     /note="synonym: PAR2"
     CDS             82. .1275
                     /gene="F2RL1"
                     /note="G-protein-coulped receptor; PAR-2"
                     /codon_start=1
                     /product="protease-activated receptor 2"
                     /protein_id="AAP97012.1"
                     /db_xref="GI:33149992"
                     /translation="MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG
                     TSHVTGKGVTVETVFSVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLF
                     RTKKKHPAVIYMANLALADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMY
                     CSILFMTCLSVQRYWVIVNPMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFI
                     PALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDE
                     NSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLST
                     LNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTT
                     VKTSY"
ORIGIN

  Query Match            88.8%;  Score 1114;  DB 9;  Length 1275;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-305;
  Matches 1114;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        139 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 221

Qy        199 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 281

Qy        259 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 341
```

```
Qy        319 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 378
Db        342 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 401

Qy        379 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 438
Db        402 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 461

Qy        439 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 498
Db        462 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 521

Qy        499 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 558
Db        522 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 581

Qy        559 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 618
Db        582 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 641

Qy        619 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 678
Db        642 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 701

Qy        679 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 738
Db        702 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 761

Qy        739 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 798
Db        762 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 821

Qy        799 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 858
Db        822 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 881

Qy        859 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 918
Db        882 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 941

Qy        919 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 978
Db        942 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 1001

Qy        979 TGTGGTGCATTATTTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1038
Db       1002 TGTGGTGCATTATTTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1061

Qy       1039 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1098
Db       1062 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1121

Qy       1099 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1158
Db       1122 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1181

Qy       1159 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1218
Db       1182 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1241

Qy       1219 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGA 1252
Db       1242 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGA 1275
```

```
RESULT 14
CQ870621
LOCUS       CQ870621               1451 bp    DNA     linear   PAT 13-SEP-2004
DEFINITION  Sequence 30 from Patent WO2004073657.
ACCESSION   CQ870621
VERSION     CQ870621.1  GI:52000132
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Aziz,N., Gish,K.C., Wilson,K.E. and Zlotnik,A.
  TITLE     Methods of diagnosis of cancer, composition and methods of
            screening for modulators of cancer
  JOURNAL   Patent: WO 2004073657-A 30 02-SEP-2004;
            PROTEIN DESIGN LABS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .1451
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             88.7%;  Score 1113.8;  DB 6;  Length 1451;
  Best Local Similarity   99.8%;  Pred. No. 6e-305;
  Matches 1115;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        139 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 198
Db        228 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 287

Qy        199 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 258
Db        288 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 347

Qy        259 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 318
Db        348 TGTCCTCACTGGAAAACTGACCACGGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 407

Qy        319 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 378
Db        408 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 467

Qy        379 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 438
Db        468 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 527

Qy        439 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 498
Db        528 GTTCCCCTTGAAGATTGCCTATCACATACATGCCAACAACTGGATTTATGGGGAAGCTCT 587

Qy        499 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 558
Db        588 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 647

Qy        559 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 618
Db        648 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 707

Qy        619 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 678
Db        708 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 767

Qy        679 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 738
Db        768 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 827

Qy        739 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 798
Db        828 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 887

Qy        799 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 858
Db        888 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 947

Qy        859 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 918
Db        948 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 1007

Qy        919 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 978
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 1067

Qy        979 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1038
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1127

Qy       1039 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1098
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1128 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1187

Qy       1099 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1158
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1188 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1247

Qy       1159 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1218
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1248 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1307

Qy       1219 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
                 |||||||||||||||||||||||||||||||||||
Db       1308 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1344


RESULT 15
CQ876755
LOCUS       CQ876755               1451 bp    DNA     linear   PAT 04-OCT-2004
DEFINITION  Sequence 1 from Patent WO2004080373.
ACCESSION   CQ876755
VERSION     CQ876755.1  GI:53790207
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Golz,S., Brueggemeier,U. and Summer,H.
  TITLE     Diagnostics and therapeutics for diseases associated with g-protein
            coupled proteinase activated receptor 2 (par2)
  JOURNAL   Patent: WO 2004080373-A 1 23-SEP-2004;
            Bayer HealthCare AG (DE)
FEATURES             Location/Qualifiers
     source          1. .1451
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             88.7%;  Score 1113.8;  DB 6;  Length 1451;
  Best Local Similarity   99.8%;  Pred. No. 6e-305;
  Matches 1115;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        139 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 198
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 287

Qy        199 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 258
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 347

Qy        259 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 318
                 |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        348 TGTCCTCACTGGAAAACTGACCACGGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 407

Qy        319 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 378
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 467

Qy        379 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 438
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 527

Qy        439 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 498
                 |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        528 GTTCCCCTTGAAGATTGCCTATCACATACATGCCAACAACTGGATTTATGGGGAAGCTCT 587

Qy        499 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 558
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        588 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 647

Qy        559 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 618
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 707

Qy        619 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 678
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        708 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 767

Qy        679 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 738
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        768 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 827

Qy        739 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 798
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        828 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 887

Qy        799 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 858
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        888 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 947

Qy        859 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 918
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        948 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 1007

Qy        919 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 978
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 1067

Qy        979 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1038
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1127

Qy       1039 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1098
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1128 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1187

Qy       1099 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1158
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1188 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1247

Qy       1159 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1218
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1248 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1307

Qy       1219 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255
                 |||||||||||||||||||||||||||||||||||
Db       1308 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1344


Search completed: March 21, 2005, 21:39:45
Job time : 5676.07 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 18, 2005, 20:59:39 ; Search time 69.0868 Seconds
(without alignments)
2228.076 Million cell updates/sec

Title: US-10-643-627-4
Perfect score: 2030
Sequence: 1 MNVLSPEQTSVTAETFISVM.....KHSRKSSSYSSSSTTVKTSY 398

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2030	100.0	398	2	AAR66921	Aar66921 Human C14
2	2030	100.0	398	2	AAW01953	Aaw01953 Human C14
3	1903	93.7	397	3	AAB35641	Aab35641 Human PAR
4	1903	93.7	397	5	AAE26678	Aae26678 Human coa
5	1903	93.7	397	6	ABG73508	Abg73508 Human par
6	1903	93.7	397	7	ADE62812	Ade62812 Human Pro
7	1903	93.7	397	8	ADO29311	Ado29311 Human GPC
8	1903	93.7	397	8	ADS74020	Ads74020 Human G-p
9	1898	93.5	397	8	ADL61221	Adl61221 Human coa
10	1897	93.4	397	6	ABP81907	Abp81907 Human pro
11	1897	93.4	397	7	ADK52594	Adk52594 Hematolog
12	1897	93.4	397	7	ADN39997	Adn39997 Cancer/an
13	1897	93.4	397	8	ADR46675	Adr46675 Cancer-as
14	1872	92.2	397	2	AAR66923	Aar66923 Human C14
15	1872	92.2	397	2	AAW01955	Aaw01955 Human C14
16	1865.5	91.9	394	2	AAW51408	Aaw51408 Human pro
17	1816	89.5	355	8	ADI28653	Adi28653 Human mod
18	1789.5	88.2	389	8	ADO28601	Ado28601 Human PAR
19	1743	85.9	341	8	ADI28654	Adi28654 Human mod
20	1706.5	84.1	395	2	AAR66920	Aar66920 Murine C1
21	1701.5	83.8	395	2	AAW01952	Aaw01952 Murine C1
22	1672	82.4	399	2	AAR66922	Aar66922 Murine C1
23	1672	82.4	399	2	AAW01954	Aaw01954 Murine C1
24	1672	82.4	399	7	ABR63562	Abr63562 Delayed h
25	1672	82.4	399	8	ADO29312	Ado29312 Mouse GPC
26	1652	81.4	397	7	ADE62810	Ade62810 Rat Prote
27	1643	80.9	320	8	ADI28655	Adi28655 Human mod
28	620	30.5	430	8	ADO29310	Ado29310 Mouse GPC
29	619	30.5	425	2	AAR27240	Aar27240 Human thr
30	619	30.5	425	2	AAR60698	Aar60698 Fragment
31	619	30.5	425	2	AAW51407	Aaw51407 Human pro
32	619	30.5	425	2	AAY49570	Aay49570 Human thr
33	619	30.5	425	5	AAE17032	Aae17032 Human thr
34	619	30.5	425	5	AAG80697	Aag80697 Human thr
35	619	30.5	425	6	ABG73511	Abg73511 Human thr
36	619	30.5	425	6	ABR47449	Abr47449 Breast ca
37	619	30.5	425	6	ABP81919	Abp81919 Human thr
38	619	30.5	425	7	ADG89876	Adg89876 Human coa
39	619	30.5	425	8	ADL14208	Adl14208 Novel hum
40	619	30.5	425	8	ADN04016	Adn04016 Antipsori
41	619	30.5	425	8	ADO29309	Ado29309 Human GPC
42	619	30.5	425	8	ADQ18985	Adq18985 Human sof
43	619	30.5	425	8	ADR45608	Adr45608 Human G p
44	619	30.5	425	8	ADS84489	Ads84489 Human pro
45	619	30.5	426	3	AAY45035	Aay45035 Human thr

ALIGNMENTS

RESULT 1
AAR66921
ID AAR66921 standard; protein; 398 AA.
XX
AC AAR66921;
XX
DT 25-MAR-2003 (revised)
DT 22-AUG-1995 (first entry)
XX
DE Human C140 receptor.
XX
KW G-protein-coupled receptor; G-protein; C140 receptor.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	Protein	1. .27
FT		/label= signal peptide
FT	Region	11. .132
FT		/label= transmembrane II
FT	Modified-site	31
FT		/label= Asn linked glycosylation site
FT	Cleavage-site	37. .38
FT		/label= protease receptor cleavage site
FT	Region	81. .103
FT		/label= transmembrane I
FT	Region	150. .174
FT		/label= transmembrane III
FT	Region	191. .212
FT		/label= transmembrane IV
FT	Modified-site	223
FT		/label= Asn linked glycosylation site
FT	Region	245. .267
FT		/label= transmembrane V
FT	Region	289. .309
FT		/label= transmembrane VI
FT	Region	327. .348
FT		/label= transmembrane VII

XX
PN WO9503318-A1.
XX
PD 02-FEB-1995.
XX
PF 26-JUL-1994; 94WO-US008536.
XX
PR 26-JUL-1993; 93US-00097938.
XX
PA (CORT-) COR THERAPEUTICS.

XX
PI Scarborough RM, Sundelin J;
XX
DR WPI; 1995-075182/10.
DR N-PSDB; AAQ84558.
XX
PT New DNA encoding recombinant C140 receptor - and novel agonists and
PT antagonists and specific antibodies with therapeutic and diagnostic
PT applications.
XX
PS Disclosure; Fig 2; 57pp; English.
XX
CC The availability of genomic DNA encoding the mouse protease C140 receptor
CC (see Q84557) permitted the retrieval of the corresp. human gene. A human
CC genomic library cloned in the vector EMBL3 was screened using the entire
CC coding region of the murine clone as a probe. The recovered human gene
CC including the DNA sequence and the deduced AA sequence are shown in
CC Q84558 & R66921. Subsequent experiments indicated that the human C140
CC gene is located in the same region of the long arm of chromosome number 5
CC (5q12-5q13) as has been reported for the human thrombin receptor gene.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 398 AA;

Query Match 100.0%; Score 2030; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 5.6e-210;
Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNVLSFEQTSVTAETFISVMTLVFLSCTGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNVLSFEQTSVTAETFISVMTLVFLSCTGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVF 60

Qy         61 SVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLAL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLAL 120

Qy        121 ADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVN 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVN 180

Qy        181 PMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDM 240

Qy        241 FNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLIC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 FNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLIC 300

Qy        301 FTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNAL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNAL 360

Qy        361 LCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||||||||||||||||
Db        361 LCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 398
```

RESULT 2
AAW01953
ID AAW01953 standard; protein; 398 AA.
XX
AC AAW01953;
XX
DT 01-APR-1997 (first entry)
XX
DE Human C140 receptor, with putative signal sequence.
XX
KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .27
FT /note= "putative signal peptide, differs from signal
FT peptide encoded by a cDNA clone of this receptor (see
FT AAW01955), the signal sequence given for the cDNA clone
FT is believed to be the correct sequence"
FT Protein 28. .398
FT /note= "mature protein"
FT Modified-site 31
FT /note= "potential Asn-linked glycosylation site"
FT Cleavage-site 37. .38
FT /note= "putative protease receptor cleavage site"
FT Region 81. .103
FT /note= "transmembrane region I"
FT Region 111. .132
FT /note= "transmembrane region II"
FT Region 151. .174
FT /note= "transmembrane region III"
FT Region 191. .212
FT /note= "transmembrane region IV"
FT Modified-site 223
FT /note= "potential Asn-linked glycosylation site"
FT Region 245. .267
FT /note= "transmembrane region V"
FT Region 289. .309
FT /note= "transmembrane region VI"
FT Region 327. .348
FT /note= "transmembrane region VII"
XX
PN WO9623225-A1.
XX
PD 01-AUG-1996.
XX
PF 25-JAN-1996; 96WO-US001179.
XX
PR 25-JAN-1995; 95US-00390301.
XX
PA (CORT-) COR THERAPEUTICS INC.
XX
PI Sundelin J, Scarborough RM;
XX
DR WPI; 1996-362813/36.
DR N-PSDB; AAT32037.
XX
PT Vector for expression C140 cell surface receptor in host cell - useful to
PT identify C140 agonist and antagonists, which are antihypertensives and
PT elevators of blood pressure, respectively.
XX
PS Example 2; Fig 2A-B; 60pp; English.
XX
CC AAW01953 represents the human C140 receptor (C140R), including a putative
CC signal peptide (see features table). DNA encoding C140R may be engineered
CC so as to allow the recombinant expression of C140R in a suitable host
CC cell, i.e. by removing the native expression-control sequences and
CC replacing them with control sequences operable in the host. Such a
CC recombinant receptor can be expressed on the surface of oocytes, this
CC provides a good assay system for identifying agonists/antagonists of
CC C140R. The C140 receptor is a G-protein linked receptor and a member of
CC the "seven-pass" transmembrane receptor superfamily (peptide chain of the
CC receptor passes through the cell membrane seven times, producing seven
CC transmembrane regions within the receptor molecule). The C140 receptor is
CC involved in controlling blood pressure. C140 antagonists (see AAW01942-
CC W01951) are useful to inhibit signalling from this receptor, resulting in
CC an increase in blood pressure and are therefore useful in pharmaceuticals
CC for the treatment of hypotension (low blood pressure). Conversely
CC agonists (see AAW01914-W01941) of C140 are useful in pharmaceuticals for
CC the treatment of hypertension (high blood pressure)
XX
SQ Sequence 398 AA;

Query Match 100.0%; Score 2030; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 5.6e-210;
Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNVLSFEQTSVTAETFISVMTLVFLSCTGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNVLSFEQTSVTAETFISVMTLVFLSCTGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVF 60

Qy         61 SVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLAL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLAL 120

Qy        121 ADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVN 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVN 180

Qy        181 PMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDM 240

Qy        241 FNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLIC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 FNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLIC 300

Qy        301 FTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNAL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNAL 360

Qy        361 LCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||||||||||||||||
Db        361 LCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 398
```

RESULT 3
AAB35641
ID AAB35641 standard; protein; 397 AA.
XX
AC AAB35641;
XX
DT 19-FEB-2001 (first entry)
XX
DE Human PAR-2 protein.
XX
KW PAR-2; protease activated receptor-2; ECL-2; inflammatory disease;
KW asthma; chronic obstructive pulmonary; arthritis; inflammatory bowel;
KW psoriasis; eczema; multiple sclerosis.
XX
OS Homo sapiens.
XX
PN WO200063371-A1.
XX
PD 26-OCT-2000.
XX
PF 17-APR-2000; 2000WO-GB001455.
XX
PR 15-APR-1999; 99GB-00008513.
XX
PA (UYSO-) UNIV SOUTHAMPTON.
XX
PI Walls AF, Palmer K, Compton SJ, Cairns JA, Gough AC;
XX
DR WPI; 2000-679599/66.
XX
PT Protease activated receptor 2 variants useful for treating inflammatory
PT diseases such as asthma, arthritis and psoriasis, and as hypertensives,
PT has reduced sensitivity to trypsin.
XX
PS Claim 2; Page 55; 59pp; English.
XX
CC The present invention relates to a variant protease activated receptor 2
CC (PAR-2). The invention is useful for identifying an individual having a
CC polymorphism in the ECL-2 region of one or both PAR-2 gene alleles. The
CC invention may be used to develop treatments for inflammatory diseases
CC such as asthma, chronic obstructive pulmonary diseases, arthritis,
CC inflammatory bowel diseases, psoriasis and eczema, multiple sclerosis and
CC to raise blood pressure
XX

SQ Sequence 397 AA;

Query Match 93.7%; Score 1903; DB 3; Length 397;
Best Local Similarity 96.9%; Pred. No. 2.9e-196;
Matches 373; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

```
Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::  |||:    ||||||||||||||||||||||||||||||||||||||||||||||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 312

Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 SLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 4
AAE26678
ID AAE26678 standard; protein; 397 AA.
XX
AC AAE26678;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human coagulation factor II (thrombin) receptor like 1 (F2RL1) protein.
XX
KW Human; haplotype; coagulation factor II receptor like 1; F2RL1; asthma;
KW polymorphism; chronic pulmonary disease; inflammatory disorder;
KW gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 30
FT /note= "This amino acid changes to Ser due to single
FT nucleotide polymorphism"
XX
PN WO200255534-A2.
XX
PD 18-JUL-2002.
XX
PF 13-NOV-2001; 2001WO-US046475.
XX
PR 10-NOV-2000; 2000US-0247516P.
XX
PA (GENA-) GENAISSANCE PHARM INC.
XX
PI Bieglecki KM, Sanchis A, Shah N;
XX
DR WPI; 2002-566728/60.
DR N-PSDB; AAD44437, AAD44438.
XX
PT New genetic variants having polymorphisms in the coagulation factor II
PT (thrombin) receptor like 1 (F2RL1) gene, useful for studying the function

PT of F2RL1 and treating disorders associated with abnormal expression or
PT function of F2RL1.
XX
PS Claim 27; Fig 3; 65pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising genes and
CC haplotypes of the coagulation factor II (thrombin) receptor like 1
CC (F2RL1) gene. Polymorphic variants of the F2RL1 gene are useful in
CC studying the expression and biological function of F2RL1, and in
CC identifying drugs targetting F2RL1 protein for treating disorders
CC associated with abnormal expression or function of F2RL1, e.g. asthma,
CC chronic pulmonary disease, and inflammatory disorders. Polynucleotides
CC comprising a polymorphic gene variant or fragment may be used for
CC therapeutic purposes, where a patient could benefit from expression or
CC increased expression of a particular F2RL1 protein isoform, or an
CC expression vector encoding the isoform may be administered to the
CC patient. Haplotype information is useful in improving the efficiency and
CC output of several steps in drug discovery and development process,
CC including target validation, identifying lead compounds, and early phase
CC clinical trials. Information on polymorphisms may be applied in studying
CC biological functions of F2RL1 as well as in identifying drugs targetting
CC this protein for the treatment of disorders related to its abnormal
CC expression or function. The invention is useful in gene therapy. The
CC present sequence is human F2RL1 protein
XX
SQ Sequence 397 AA;

Query Match 93.7%; Score 1903; DB 5; Length 397;
Best Local Similarity 96.9%; Pred. No. 2.9e-196;
Matches 373; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

```
Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::    |||:    ||||||||||||||||||||||||||||||||||||||||||||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 312

Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 SLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 5
ABG73508
ID ABG73508 standard; protein; 397 AA.
XX
AC ABG73508;
XX
DT 14-FEB-2003 (first entry)
XX
DE Human par2 protein SEQ ID 39.
XX
KW G-protein coupled receptor; HGPRBMY1; HGPRBMY2; immunosuppressive;
KW cardiant; neuroprotective; antiinflammatory; cytostatic; vulnerary;
KW vaccine; gene therapy; autoimmune; cardiovascular; neural; reproductive;
KW haematopoietic; pulmonary; gastrointestinal; proliferation; cell cycle;
KW birth defect; aberrant phosphorylation; acute phase response; receptor;
KW signal transduction; hyperimmune activity; inflammatory; hypercongenital;
KW necrotic lesion; wound; organ transplant rejection.
XX
OS Homo sapiens.
XX
PN WO200268591-A2.
XX
PD 06-SEP-2002.
XX
PF 22-FEB-2002; 2002WO-US005281.
XX
PR 23-FEB-2001; 2001US-0270792P.
PR 23-FEB-2001; 2001US-0270793P.
PR 06-JUN-2001; 2001US-0296427P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Feder J, Ramanathan C, Nelson T, Mintier G, Cacace A, Barber L;
PI Kornacker M, Bol D;
XX
DR WPI; 2003-058304/05.
XX
PT New human HGPRBMY1 or HGPRBMY2 polynucleotide and polypeptide, useful
PT preventing, treating or ameliorating a disorder e.g., wound,
PT cardiovascular disorder or transplant rejection.
XX
PS Disclosure; Fig 4; 316pp; English.
XX
CC This invention describes the novel human G-protein coupled receptors
CC (GPCR's), HGPRBMY1 or HGPRBMY2 which have immunosuppressive, cardiant,
CC neuroprotective, antiinflammatory, cytostatic and vulnerary activity and
CC can be used in vaccines or for gene therapy. Pharmaceutical compositions
CC comprising HGPRBMY1 or HGPRBMY2 polypeptides or their agonists or
CC antagonists or modulators, or a HGPRBMY1- or HGPRBMY2-specific antibody
CC are useful for preventing, treating or ameliorating a medical condition
CC comprising autoimmune, cardiovascular, neural, reproductive,
CC haematopoietic, pulmonary, gastrointestinal or proliferating disorder, a
CC cell cycle or birth defect, a disorder related to aberrant
CC phosphorylation, acute phase responses or signal transduction or to
CC hyperimmune activity, an inflammatory or hypercongenital condition, a
CC necrotic lesion, a wound, organ transplant rejection or a condition
CC related to organ transplant rejection. This sequence represents a G-
CC protein coupled receptor associated with the human HGPRBMY proteins
CC described in the disclosure of the invention
XX
SQ Sequence 397 AA;

Query Match 93.7%; Score 1903; DB 6; Length 397;
Best Local Similarity 96.9%; Pred. No. 2.9e-196;
Matches 373; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

```
Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::    |||:    ||||||||||||||||||||||||||||||||||||||||||||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 312
```

```
Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 SLTSKKHSRKSSSYSSSSTTVKTSY 397


RESULT 6
ADE62812
ID   ADE62812 standard; protein; 397 AA.
XX
AC   ADE62812;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human Protein P55085, SEQ ID NO 8745.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; P55085.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a human protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 397 AA;

  Query Match             93.7%;  Score 1903;  DB 7;  Length 397;
  Best Local Similarity   96.9%;  Pred. No. 2.9e-196;
  Matches  373;  Conservative    4;  Mismatches    4;  Indels    4;  Gaps    1;

Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::    |||:    ||||||||||||||||||||||||||||||||||||||||||||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 312

Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 SLTSKKHSRKSSSYSSSSTTVKTSY 397


RESULT 7
ADO29311
ID   ADO29311 standard; protein; 397 AA.
XX
AC   ADO29311;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human GPCR F2RL1, SEQ ID NO:412.
XX
KW   G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW   transgenic mouse; neurological disorder; adrenal gland disorder;
KW   colon disorder; intestinal disorder; cardiovascular disorder;
KW   muscular disorder; blood disorder; immune disorder; bone disorder;
KW   joint disorder; metabolic disorder; nutritive disorder; cancer;
KW   kidney disorder; liver disorder; lung disorder; breast disorder;
KW   ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW   skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW   thymus disorder; thyroid disorder; antiparkinsonian; antimanic;
KW   cytostatic; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
KW   CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW   virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW   dermatological; antiulcer; antithyroid; antiallergic; anorectic;
KW   immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW   receptor.
XX
OS   Homo sapiens.
XX
PN   WO2004040000-A2.
XX
PD   13-MAY-2004.
XX
PF   09-SEP-2003; 2003WO-US028226.
XX
PR   09-SEP-2002; 2002US-0409303P.
```

```
PR   09-APR-2003; 2003US-0461329P.
XX
PA   (PRIM-) PRIMAL INC.
XX
PI   Gaitanaris GA,  Bergmann JE,  Gragerov A,  Hohmann J,  Li F;
PI   Madisen L,  Mcilwain KL,  Pavlova MN,  Vassilatis D,  Zeng H;
XX
DR   WPI; 2004-390329/36.
DR   N-PSDB; ADO29874.
XX
PT   Novel mammalian G protein coupled receptors, useful for identifying
PT   compounds that modulates diagnosing and treating disease condition
PT   associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT   pectoris, Parkinson's disease.
XX
PS   Claim 151; SEQ ID NO 412; 542pp; English.
XX
CC   The invention relates to human and mouse G protein-coupled receptors
CC   (GPCRs) and nucleic acids encoding them. The invention also relates to
CC   sequences at least 90% identical to the GPCR proteins and nucleic acids
CC   of the invention; methods of treating, preventing or diagnosing diseases
CC   associated with GPCRs of the invention; methods of screening for
CC   compounds useful in the treatment of GPCR-related diseases; a transgenic
CC   mouse comprising a GPCR gene of the invention; a mouse comprising a
CC   mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC   from the trasngenic mice; kits comprising several mice, each of which has
CC   a mutation in a different GPCR gene of the invention; and kits comprising
CC   probes which hybridise to GPCR polynucleotides of the invention. The
CC   invention further discloses variants of the GPCR polypeptides and vectors
CC   comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC   be used in the diagnosis, treatment or prevention of a wide variety of
CC   diseases including neurological disorders (e.g., Alzheimer's disease,
CC   depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC   disorders of the adrenal gland; disorders of the colon or intestine
CC   (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC   syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC   myocardial infarction); muscular disorders; blood disorders (e.g.,
CC   anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC   AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC   arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC   obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC   diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC   uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC   thyroid (e.g., cancers). The present sequence represents a GPCR of the
CC   invention. Note: The full sequence data for this patent did not form part
CC   of the printed specification; those sequences not shown were obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 397 AA;

  Query Match             93.7%;  Score 1903;  DB 8;  Length 397;
  Best Local Similarity   96.9%;  Pred. No. 2.9e-196;
  Matches  373;  Conservative    4;  Mismatches    4;  Indels    4;  Gaps    1;

Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::    |||:    ||||||||||||||||||||||||||||||||||||||||||||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 312

Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 SLTSKKHSRKSSSYSSSSTTVKTSY 397


RESULT 8
ADS74020
ID   ADS74020 standard; protein; 397 AA.
XX
AC   ADS74020;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Human G-protein coupled proteinase activated receptor 2 (PAR2).
XX
KW   Human; proteinase activated receptor 2; PAR2; G-protein coupled receptor;
KW   receptor; cardiant; neuroprotective; nephrotropic; respiratory-gen.;
KW   gastrointestinal-gen.; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO2004080373-A2.
XX
PD   23-SEP-2004.
XX
PF   26-FEB-2004; 2004WO-EP001896.
XX
PR   11-MAR-2003; 2003EP-00004980.
XX
PA   (FARB ) BAYER HEALTHCARE AG.
XX
PI   Golz S,  Brueggemeier U,  Summer H;
XX
DR   WPI; 2004-677358/66.
DR   N-PSDB; ADS74019.
XX
PT   Screening for therapeutic agents for treating e.g., cardiovascular
PT   diseases by contacting a test compound with a proteinase activated
PT   receptor 2 (PAR2) polypeptide or polynucleotide and detecting binding of
PT   the test compound.
XX
PS   Disclosure; SEQ ID NO 2; 121pp; English.
XX
CC   The present sequence is that of human G-protein coupled proteinase
CC   activated receptor 2 (PAR2). PAR2 is an antiinflammatory receptor in the
CC   colon and may also play a role in the airway, regulating sodium ion
CC   absorption and anion secretion. The invention relates to novel disease
CC   associations of PAR2 polypeptides and polynucleotides. It also relates to
CC   novel methods of screening for therapeutic agents for the treatment of
CC   cardiovascular disorders, gastrointestinal and liver diseases,
CC   neurological disorders, urological disorders, haematological diseases and
CC   respiratory diseases in a mammal. Suitable therapeutic agents include a
CC   small molecule, an RNA molecule, an antisense oligonucleotide, a
CC   polypeptide, an antibody or a ribozyme. The invention also provides
CC   pharmaceutical compositions for the treatment of diseases and disorders
CC   associated with PAR2 comprising a PAR2 polypeptide, PAR2 polynucleotide
CC   or a regulator or modulator of PAR2 activity. Methods of diagnosing these
CC   diseases and disorders involve determining the amount of PAR2
CC   polynucleotide in a sample.
XX
SQ   Sequence 397 AA;

  Query Match             93.7%;  Score 1903;  DB 8;  Length 397;
  Best Local Similarity   96.9%;  Pred. No. 2.9e-196;
  Matches  373;  Conservative    4;  Mismatches    4;  Indels    4;  Gaps    1;

Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
```

```
           :::    |||:    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 312

Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
           |||||||||||||||||||||||||
Db        373 SLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 9
ADL61221
ID ADL61221 standard; protein; 397 AA.
XX
AC ADL61221;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human coagulation factor II (thrombin) receptor-like 1 protein.
XX
KW predictor set; protein tyrosine kinase; cytostatic; antiangiogenic;
KW vasotropic; vulnerary; pharmacogenomic; drug sensitivity; breast cancer;
KW hypervascular disease; angiogenesis; wound healing scar; human;
KW biomarker; coagulation factor II receptor-like 1; thrombin; receptor.
XX
OS Homo sapiens.
XX
PN WO2004020583-A2.
XX
PD 11-MAR-2004.
XX
PF 26-AUG-2003; 2003WO-US026491.
XX
PR 27-AUG-2002; 2002US-0406385P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Huang F, Han X, Reeves KA, Amler L, Fairchild CR, Lee FY;
PI Shaw P;
XX
DR WPI; 2004-239171/22.
DR N-PSDB; ADL61084.
XX
PT New predictor sets with a plurality of polynucleotides and/or
PT polypeptides whose expression pattern predicts cell response to a
PT compound that modulates protein tyrosine kinase activity, useful in
PT treating breast cancer.
XX
PS Claim 9; SEQ ID NO 145; 649pp; English.
XX
CC The invention relates to a novel predictor set comprising a plurality of
CC polynucleotides and/or polypeptides whose expression pattern is
CC predictive of the response of cells to treatment with a compound that
CC modulates protein tyrosine kinase activity or members of the protein
CC tyrosine kinase pathway. The molecules of the invention demonstrate
CC cytostatic, antiangiogenic, vasotropic and vulnerary activities and may
CC be useful in the field of pharmacogenomics, in particular for determining
CC drug sensitivity and in treating breast cancer, hypervascular diseases,
CC angiogenesis and scars in wound healing. The current sequence is that of
CC a human protein tyrosine kinase biomarker protein of the invention.
XX
SQ Sequence 397 AA;

```
  Query Match           93.5%;  Score 1898;  DB 8;  Length 397;
  Best Local Similarity  96.6%;  Pred. No. 9.9e-196;
  Matches  372;  Conservative    5;  Mismatches    4;  Indels    4;  Gaps    1;

Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::    |||:    ||:||||||||||||||||||||||||||||||||||||||||||
Db         13 AILLAASLSCSGTIQGTSRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 312

Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 SLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 10
ABP81907
ID ABP81907 standard; protein; 397 AA.
XX
AC ABP81907;
XX
DT 04-MAR-2003 (first entry)
XX
DE Human proteinase-activated receptor 2 protein SEQ ID NO:300.
XX
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer.
XX
OS Homo sapiens.
XX
PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US050107.
XX
PR 19-DEC-2000; 2000US-0257144P.
XX

```
PA   (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI   Burmer GC,  Roush CL,  Brown JP;
XX
DR   WPI; 2003-046718/04.
DR   N-PSDB; ABZ42755.
XX
PT   New isolated antigenic peptides e.g., for G protein-coupled receptors
PT   (GPCR), useful for diagnosing and designing drugs for treating conditions
PT   in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT   autoimmune diseases.
XX
PS   Disclosure; Fig 1; 523pp; English.
XX
CC   The present invention describes antigenic peptides (I) comprising: (a)
CC   any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC   acids. Also described: (1) an assay for the detection of a particular G
CC   protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC   and (2) an isolated antibody having high specificity and high affinity or
CC   avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC   gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC   antibody against a particular GPCR, and in the production of specific
CC   antibodies. The peptides and antibodies are also useful for detecting the
CC   presence or absence of corresponding GPCRs. The antigenic peptides for
CC   GPCRs and antibodies are useful for diagnosing and designing drugs for
CC   treating immune-related diseases, growth-related diseases, cell
CC   regeneration-related disease, immunological-related cell proliferative
CC   diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC   atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC   osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC   inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC   disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC   anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC   loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC   hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC   any other disorder in which GPCRs are involved. The antibodies may be
CC   used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC   GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC   exemplification of the present invention
XX
SQ   Sequence 397 AA;

  Query Match             93.4%;  Score 1897;  DB 6;  Length 397;
  Best Local Similarity   96.6%;  Pred. No. 1.3e-195;
  Matches  372;  Conservative    4;  Mismatches    5;  Indels    4;  Gaps    1;

Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::    |||:    ||||||||||||||||||||||||||||||||||||||||||||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHANNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 312

Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 SLTSKKHSRKSSSYSSSSTTVKTSY 397
```

```
RESULT 11
ADK52594
ID   ADK52594 standard; protein; 397 AA.
XX
AC   ADK52594;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Hematological disorder associated Gene ID 340 encoded protein.
XX
KW   cytostatic; antianemic; antisickling; virucide; hemostatic; nephrotropic;
KW   cytostatic; thrombolytic; antiparasitic; gene therapy;
KW   hematologic disorder; cancer; Sickle Cell Anemia;
KW   Infectious Mononucleosis; Leukemia; Polycythemia Vera; Lymphoma;
KW   Retinoblastoma; Hemophilia; Thrombosis; Herpes; Thalassemia;
KW   transfusion reaction; Erythroblastosis; mechanical trauma;
KW   micro-angiopathic hemolytic anemia; parasite infection.
XX
OS   Homo sapiens.
XX
PN   WO2003065871-A2.
XX
PD   14-AUG-2003.
XX
PF   28-JAN-2003; 2003WO-US002484.
XX
PR   04-FEB-2002; 2002US-0354333P.
PR   28-FEB-2002; 2002US-0360258P.
PR   15-MAR-2002; 2002US-0364476P.
PR   26-APR-2002; 2002US-0375626P.
PR   06-JUN-2002; 2002US-0386494P.
PR   24-JUN-2002; 2002US-0390965P.
PR   28-JUN-2002; 2002US-0392480P.
PR   03-JUL-2002; 2002US-0394128P.
PR   31-JUL-2002; 2002US-0399783P.
PR   13-AUG-2002; 2002US-0403221P.
PR   30-AUG-2002; 2002US-0407045P.
PR   25-NOV-2002; 2002US-0429048P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Carroll JM,  Healy A,  Weich NS,  Kelly LM;
XX
DR   WPI; 2003-731464/69.
DR   N-PSDB; ADK52593.
XX
PT   Identifying a compound capable of treating a hematologic disorder (e.g.
PT   anemia or leukemia) comprises assaying the ability of the compound to
PT   modulate the expression or activity of e.g. 131,148, 199 or 12303
PT   polypeptide or nucleic acid.
XX
PS   Disclosure; SEQ ID NO 52; 232pp; English.
XX
CC   The invention relates to a method of identifying a compound capable of
CC   treating a hematologic disorder comprises assaying the ability of the
CC   compound to modulate 131,148, 199, 12303, 13906, 15513, 17822, 302, 5677,
CC   194, 14393, 28059, 7366, 12212, 1981, 261, 12416, 270, 1410, 137, 1871,
CC   13051, 1847, 1849, 15402, 340, 10217, 837, 1761, 8990 or 13249 nucleic
CC   acid expression or polypeptide activity, thus, identifying a compound
CC   capable of treating a hematologic disorder. The methods are useful in
CC   diagnosing, preventing and treating hematological disorders, such as
CC   cancer, Sickle Cell Anemia, Infectious Mononucleosis, Leukemia,
CC   Polycythemia Vera, Lymphoma, Retinoblastoma, Hemophilia, disorders
CC   associated with an increased risk of Thrombosis, Herpes, Thalassemia,
CC   antibody-mediated disorders such as transfusion reactions and
CC   Erythroblastosis, mechanical trauma to red blood cells such as micro-
CC   angiopathic hemolytic anemias, infections by parasites or chemical
CC   injuries. The methods may also be used for identifying compounds that
CC   modulate hematological disorders. This sequence corresponds to the
CC   protein encoded by one of the genes modulated by the compounds.
XX
```

```
SQ   Sequence 397 AA;

  Query Match             93.4%;  Score 1897;  DB 7;  Length 397;
  Best Local Similarity   96.6%;  Pred. No. 1.3e-195;
  Matches  372;  Conservative    4;  Mismatches    5;  Indels    4;  Gaps    1;

Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::    |||:    |||||||||||||||||||||||||||||||||||||||||||||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHANNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 312

Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 SLTSKKHSRKSSSYSSSSTTVKTSY 397


RESULT 12
ADN39997
ID   ADN39997 standard; protein; 397 AA.
XX
AC   ADN39997;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C367.
XX
KW   Human; differential expression; cancer; angiogenic disorder;
KW   fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW   inflammatory disease; autoimmune disease;
KW   retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW   detection; diagnosis; prognosis; drug screening; drug targeting;
KW   wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW   vulnerary; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO2003042661-A2.
XX
PD   22-MAY-2003.
XX
PF   13-NOV-2002; 2002WO-US036810.
XX
PR   13-NOV-2001; 2001US-0350666P.
PR   21-NOV-2001; 2001US-0332464P.
PR   29-NOV-2001; 2001US-0334393P.
PR   03-DEC-2001; 2001US-0335394P.
PR   14-DEC-2001; 2001US-0340376P.
PR   08-JAN-2002; 2002US-0347211P.
PR   10-JAN-2002; 2002US-0347349P.
PR   08-FEB-2002; 2002US-0355250P.
PR   13-FEB-2002; 2002US-0356714P.
PR   20-FEB-2002; 2002US-0359077P.
PR   29-MAR-2002; 2002US-0368809P.
PR   04-APR-2002; 2002US-0370110P.
PR   12-APR-2002; 2002US-0372246P.
PR   05-JUN-2002; 2002US-0386614P.
PR   16-JUL-2002; 2002US-0396839P.
PR   22-JUL-2002; 2002US-0397775P.
PR   22-JUL-2002; 2002US-0397845P.
PR   09-SEP-2002; 2002US-0409450P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Afar D,  Aziz N,  Ginsburg WM,  Gish KC,  Glynne R,  Hevezi PA;
PI   Mack DH,  Murray R,  Watson SR,  Wilson KE,  Zlotnik A;
XX
DR   WPI; 2003-468649/44.
DR   N-PSDB; ADN39780.
XX
PT   Determining the presence or absence of a pathological cell in a patient,
PT   useful for diagnosing, prognosing or treating cancer, comprises detecting
PT   a nucleic acid in a biological sample.
XX
PS   Claim 12; SEQ ID NO C367; 1385pp; English.
XX
CC   The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC   whose expression is upregulated or downregulated in specific cancers or
CC   other diseases such as angiogenic or fibrotic disorders, and to methods
CC   of determining the presence or absence of a pathological cell in a
CC   patient by detecting a nucleic acid at least 80% identical to those of
CC   the invention or by detecting a polypeptide of the invention. The
CC   invention also relates to expression vectors and host cells comprising a
CC   nucleic acid of the invention; antibodies which specifically bind a
CC   polypeptide of the invention; use of such antibodies for drug targeting;
CC   and methods of screening for modulators of activity or expression of the
CC   polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC   antibodies and methods are useful for diagnosing, prognosing and treating
CC   cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC   atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC   neovascularistaion syndromes, scarring and uterine fibroids. They may
CC   also be useful in wound healing and in contraception. The present
CC   sequence represents a polypeptide of the invention.
XX
SQ   Sequence 397 AA;

  Query Match             93.4%;  Score 1897;  DB 7;  Length 397;
  Best Local Similarity   96.6%;  Pred. No. 1.3e-195;
  Matches  372;  Conservative    4;  Mismatches    5;  Indels    4;  Gaps    1;

Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::    |||:    |||||||||||||||||||||||||||||||||||||||||||||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHANNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 312

Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 SLTSKKHSRKSSSYSSSSTTVKTSY 397
```

```
RESULT 13
ADR46675
ID   ADR46675 standard; protein; 397 AA.
XX
AC   ADR46675;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Cancer-associated protein, SEQ ID 88.
XX
KW   Cytostatic; Gene Therapy; cancer; human.
XX
OS   Homo sapiens.
XX
PN   WO2004073657-A2.
XX
PD   02-SEP-2004.
XX
PF   19-FEB-2004; 2004WO-US005455.
XX
PR   19-FEB-2003; 2003US-0448784P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Gish KC,  Wilson KE,  Zlotnik A;
XX
DR   WPI; 2004-652787/63.
DR   N-PSDB; ADR46617.
XX
PT   Detecting a pathological cell in a patient for diagnosing or treating
PT   cancer by detecting in a biological sample from the patient genes whose
PT   expression are up-regulated or down-regulated in specific cancers.
XX
PS   Claim 1; SEQ ID NO 88; 375pp; English.
XX
CC   The present invention relates to a method for detecting cancer in a
CC   patient. The method comprises detecting in a biological sample from the
CC   patient a nucleotide or protein sequence comprising a sequence that is at
CC   least 80% identical to a nucleotide sequence (ADR46588-ADR46645) or
CC   protein sequence (ADR46646-ADR46703). The method is useful for detecting
CC   cancer for preparing a composition for diagnosing or treating cancer.
XX
SQ   Sequence 397 AA;

  Query Match             93.4%;  Score 1897;  DB 8;  Length 397;
  Best Local Similarity   96.6%;  Pred. No. 1.3e-195;
  Matches  372;  Conservative    4;  Mismatches    5;  Indels    4;  Gaps    1;

Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::    |||:    |||||||||||||||||||||||||||||||||||||||||||||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHANNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 312

Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 SLTSKKHSRKSSSYSSSSTTVKTSY 397


RESULT 14
AAR66923
ID   AAR66923 standard; protein; 397 AA.
XX
AC   AAR66923;
XX
DT   25-MAR-2003  (revised)
DT   22-AUG-1995  (first entry)
XX
DE   Human C140 receptor encoded by cDNA.
XX
KW   G-protein-coupled receptor; G-protein; C140 receptor.
XX
OS   Homo sapiens.
XX
PN   WO9503318-A1.
XX
PD   02-FEB-1995.
XX
PF   26-JUL-1994;   94WO-US008536.
XX
PR   26-JUL-1993;   93US-00097938.
XX
PA   (CORT-) COR THERAPEUTICS.
XX
PI   Scarborough RM,  Sundelin J;
XX
DR   WPI; 1995-075182/10.
DR   N-PSDB; AAQ84560.
XX
PT   New DNA encoding recombinant C140 receptor - and novel agonists and
PT   antagonists and specific antibodies with therapeutic and diagnostic
PT   applications.
XX
PS   Example; Fig 11; 57pp; English.
XX
CC   A human intestinal tumour cDNA library was subjected to PCR using primers
CC   designed from the genomic clone (see AAQ84558) and the amplified fragment
CC   was cloned in pSG5 and sequenced. There are four AA differences between
CC   the cDNA encoded sequence and that encoded by the genomic DNA. The
CC   genomic DNA sequence and deduced AA sequence are given in AAQ84560 &
CC   AAR66923. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 397 AA;

  Query Match             92.2%;  Score 1872;  DB 2;  Length 397;
  Best Local Similarity   95.3%;  Pred. No. 6.3e-193;
  Matches  367;  Conservative    4;  Mismatches   10;  Indels    4;  Gaps    1;

Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::    |||:    |||||||||||||||||||||||||||||||||||||||||| ||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLAGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        133 AYHIHGNNWIYGEALCNVLIGFFYRNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICFTPSNLLLVVHYF 312
```

```
Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 PLTSKKHSRKSSSYSSSSTTVKTSY 397


RESULT 15
AAW01955
ID   AAW01955 standard; protein; 397 AA.
XX
AC   AAW01955;
XX
DT   02-APR-1997  (first entry)
XX
DE   Human C140 receptor.
XX
KW   C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW   antagonist; hypertension; hypotension; blood pressure.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .27
FT                   /note= "the signal peptide differs from that encoded by a
FT                   genomic DNA sequence for this receptor (see AAW01953),
FT                   the signal sequence given here is believed to be the
FT                   correct sequence"
FT   Protein         28. .397
FT                   /note= "mature protein"
XX
PN   WO9623225-A1.
XX
PD   01-AUG-1996.
XX
PF   25-JAN-1996;   96WO-US001179.
XX
PR   25-JAN-1995;   95US-00390301.
XX
PA   (CORT-) COR THERAPEUTICS INC.
XX
PI   Sundelin J,  Scarborough RM;
XX
DR   WPI; 1996-362813/36.
DR   N-PSDB; AAT32039.
XX
PT   Vector for expression C140 cell surface receptor in host cell - useful to
PT   identify C140 agonist and antagonists, which are antihypertensives and
PT   elevators of blood pressure, respectively.
XX
PS   Example 5; Fig 11A-B; 60pp; English.
XX
CC   AAW01955 represents the human C140 receptor (C140R). DNA encoding C140R
CC   may be engineered so as to allow the recombinant expression of C140R in a
CC   suitable host cell, i.e. by removing the native expression-control
CC   sequences and replacing them with control sequences operable in the host.
CC   Such a recombinant receptor can be expressed on the surface of oocytes,
CC   this provides a good assay system for identifying agonists/antagonists of
CC   C140R. The C140 receptor is a G-protein linked receptor and a member of
CC   the "seven-pass" transmembrane receptor superfamily (peptide chain of the
CC   receptor passes through the cell membrane seven times, producing seven
CC   transmembrane regions within the receptor molecule). The C140 receptor is
CC   involved in controlling blood pressure. C140 antagonists (see AAW01942-
CC   W01951) are useful to inhibit signalling from this receptor, resulting in
CC   an increase in blood pressure and are therefore useful in pharmaceuticals
CC   for the treatment of hypotension (low blood pressure). Conversely
CC   agonists (see AAW01914-W01941) of C140 are useful in pharmaceuticals for
CC   the treatment of hypertension (high blood pressure)
XX
SQ   Sequence 397 AA;
```

```
  Query Match           92.2%;  Score 1872;  DB 2;  Length 397;
  Best Local Similarity  95.3%;  Pred. No. 6.3e-193;
  Matches  367;  Conservative    4;  Mismatches   10;  Indels    4;  Gaps    1;

Qy         18 SVMTLVFLSCT----GTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK 73
              :::   |||:    ||||||||||||||||||||||||||||||||||||||||||| ||
Db         13 AILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLAGK 72

Qy         74 LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 133
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         73 LTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPLKI 132

Qy        134 AYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 193
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        133 AYHIHGNNWIYGEALCNVLIGFFYRNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI 192

Qy        194 GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 253
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFL 252

Qy        254 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYF 313
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        253 FPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICFTPSNLLLVVHYF 312

Qy        314 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 LIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQV 372

Qy        374 SLTSKKHSRKSSSYSSSSTTVKTSY 398
              |||||||||||||||||||||||||
Db        373 PLTSKKHSRKSSSYSSSSTTVKTSY 397


Search completed: March 18, 2005, 21:06:43
Job time : 73.0868 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 21, 2005, 12:11:20 ; Search time 825.407 Seconds
(without alignments)
10141.072 Million cell updates/sec

Title: US-10-643-627-62
Perfect score: 1414
Sequence: 1 caaagaattgtaatacgact.....acaaatttcacaaataaagc 1414

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1414	100.0	1414	2	AAQ84560	Aaq84560 Human C14
2	1414	100.0	1414	2	AAT32039	Aat32039 Human C14
3	1279.4	90.5	2876	12	ADO28600	Ado28600 Human PAR
4	1273	90.0	1451	3	AAC60319	Aac60319 Human PAR
5	1273	90.0	1451	8	ABZ42755	Abz42755 Human pro
6	1273	90.0	1451	10	ADK52593	Adk52593 Hematolog
7	1273	90.0	1451	11	ADN39780	Adn39780 Cancer/an
8	1273	90.0	1451	13	ADR46617	Adr46617 Cancer-as
9	1273	90.0	1451	13	ADS74019	Ads74019 Human G-p
10	1263.6	89.4	2848	6	ABZ35045	Abz35045 Human gen
11	1188.4	84.0	8624	6	AAD44437	Aad44437 Human coa
12	1182.8	83.6	1194	6	AAD44438	Aad44438 Human coa
13	1182.8	83.6	1289	3	AAZ50773	Aaz50773 Human pro
14	1179.6	83.4	1194	12	ADO29874	Ado29874 Human GPC
15	1105.8	78.2	1255	2	AAQ84558	Aaq84558 Human C14
16	1105.8	78.2	1255	2	AAT32037	Aat32037 Human C14
17	1055.2	74.6	1068	12	ADI28650	Adi28650 Human mod
18	1013.2	71.7	1026	12	ADI28651	Adi28651 Human mod
19	951.8	67.3	963	12	ADI28652	Adi28652 Human mod
20	833.8	59.0	2713	9	ACC85274	Acc85274 Delayed h
21	833.8	59.0	2732	2	AAQ84559	Aaq84559 Murine C1
22	833.8	59.0	2732	2	AAT32038	Aat32038 Murine C1
23	829.2	58.6	1200	12	ADO30165	Ado30165 Mouse GPC
24	817.8	57.8	1428	10	ABT42343	Abt42343 Toxicity
25	805.4	57.0	1477	2	AAQ84557	Aaq84557 Murine C1
26	803.8	56.8	1477	2	AAT32036	Aat32036 Murine C1
27	383	27.1	486	6	ABK55138	Abk55138 Human col
28	303.2	21.4	493	6	ABK27602	Abk27602 Human col
29	303.2	21.4	641	6	ABL37070	Abl37070 Human col
30	166.4	11.8	3418	10	ABT41875	Abt41875 Toxicity
31	160.6	11.4	1116	6	ABK70888	Abk70888 Human cDN
32	160.6	11.4	1209	6	ABK70887	Abk70887 Human cDN
33	160.6	11.4	1278	6	ABK70889	Abk70889 Human cDN
34	160.6	11.4	3590	13	ACN38232	Acn38232 Tumour-as
35	156.6	11.1	6203	10	ADG89941	Adg89941 Human coa
36	155.8	11.0	1278	10	ADG89942	Adg89942 Human coa
37	155.8	11.0	1278	12	ADO29873	Ado29873 Human GPC
38	155.8	11.0	1764	2	AAQ73590	Aaq73590 Fragment
39	155.8	11.0	2910	2	AAT62461	Aat62461 ThR-GPA1
40	155.8	11.0	3182	3	AAA35310	Aaa35310 Human ade
41	155.8	11.0	3182	3	AAF21432	Aaf21432 Human low
42	155.8	11.0	3182	10	ABZ97126	Abz97126 Human nuc
43	155.8	11.0	3182	11	ABD20975	Abd20975 Human pul
44	155.8	11.0	3299	9	ADA24508	Ada24508 Human cDN
45	155.8	11.0	3472	2	AAZ32191	Aaz32191 Human thr

ALIGNMENTS

RESULT 1
AAQ84560
ID AAQ84560 standard; cDNA; 1414 BP.
XX
AC AAQ84560;
XX
DT 25-MAR-2003 (revised)
DT 22-AUG-1995 (first entry)
XX
DE Human C140 receptor cDNA.
XX
KW G-protein-coupled receptor; G-protein; C140 receptor; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 50. .1243
FT /*tag= a
XX
PN WO9503318-A1.
XX
PD 02-FEB-1995.
XX
PF 26-JUL-1994; 94WO-US008536.
XX
PR 26-JUL-1993; 93US-00097938.
XX
PA (CORT-) COR THERAPEUTICS.
XX
PI Scarborough RM, Sundelin J;
XX
DR WPI; 1995-075182/10.
DR P-PSDB; AAR66923.
XX
PT New DNA encoding recombinant C140 receptor - and novel agonists and
PT antagonists and specific antibodies with therapeutic and diagnostic
PT applications.
XX
PS Claim 1; Fig 11; 57pp; English.
XX
CC A human intestinal tumour cDNA library was subjected to PCR using primers
CC designed from the genomic clone (see AAQ84558) and the amplified fragment
CC was cloned in pSG5 and sequenced. There are four AA differences between

```
CC   the cDNA encoded sequence and that encoded by the genomic DNA. The
CC   genomic DNA sequence and deduced AA sequence are given in AAQ84560 &
CC   AAR66923. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 1414 BP; 335 A; 361 C; 309 G; 409 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 1414;  DB 2;  Length 1414;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1414;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CAAAGAATTGTAATACGACTCACTATAGGGCGAATTCGGATCCAGGAGGATGCGGAGCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAAGAATTGTAATACGACTCACTATAGGGCGAATTCGGATCCAGGAGGATGCGGAGCCC 60

Qy         61 CAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGG 120

Qy        121 CACCATCCAAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACCATCCAAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGG 180

Qy        181 CACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTT 240

Qy        241 TTCTGCATCTGTCCTCGCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCTGCATCTGTCCTCGCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAAT 300

Qy        301 TGTGTTTGCGGTGGGTTTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGTGTTTGCGGTGGGTTTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAAC 360

Qy        361 TAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTC 420

Qy        421 TGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGG 480

Qy        481 GGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCT 540

Qy        541 CTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTC 600

Qy        601 CAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCT 660

Qy        661 GGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCAC 720

Qy        721 GACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCT 780

Qy        781 CTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCT 840

Qy        841 GATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAG 900

Qy        901 GGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACTCCTAGTAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACTCCTAGTAA 960

Qy        961 CCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGC 1020

Qy       1021 CCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTA 1080

Qy       1081 TTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGT 1140

Qy       1141 CCGCACTGTAAAGCAGATGCAAGTACCCCTCACCTCAAAGAAACACTCCAGGAAATCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCGCACTGTAAAGCAGATGCAAGTACCCCTCACCTCAAAGAAACACTCCAGGAAATCCAG 1200

Qy       1201 CTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTTTTCCAGGTCCTCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTTTTCCAGGTCCTCAG 1260

Qy       1261 ATGGGAATTGCACAGTAGGATGTGGAACCTGTTTAATGTTATGAGGACGTGTCTGTTATT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATGGGAATTGCACAGTAGGATGTGGAACCTGTTTAATGTTATGAGGACGTGTCTGTTATT 1320

Qy       1321 TCCGGATCCAGATCTTATTAAAGCAGAACTTGTTTATTGCAGCTTATAATGGTTACAAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TCCGGATCCAGATCTTATTAAAGCAGAACTTGTTTATTGCAGCTTATAATGGTTACAAAT 1380

Qy       1381 AAAGCAATAGCATCACAAATTTCACAAATAAAGC 1414
              |||||||||||||||||||||||||||||||||
Db       1381 AAAGCAATAGCATCACAAATTTCACAAATAAAGC 1414
```

```
RESULT 2
AAT32039
ID   AAT32039 standard; cDNA; 1414 BP.
XX
AC   AAT32039;
XX
DT   02-APR-1997  (first entry)
XX
DE   Human C140 receptor cDNA clone.
XX
KW   C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW   antagonist; hypertension; hypotension; blood pressure; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             50. .1243
FT                   /*tag=  a
FT   sig_peptide     50. .130
FT                   /*tag=  b
FT                   /note= "the signal sequence, differs from the signal
FT                   sequence shown in a genomic clone of this receptor (see
FT                   AAT32037). The cDNA clone also encodes a mature receptor
FT                   having four amino acid changes, when compared to the
FT                   genomic clone"
FT   mat_peptide     131. .1240
FT                   /*tag=  c
XX
PN   WO9623225-A1.
XX
PD   01-AUG-1996.
XX
PF   25-JAN-1996;   96WO-US001179.
XX
PR   25-JAN-1995;   95US-00390301.
XX
PA   (CORT-) COR THERAPEUTICS INC.
XX
PI   Sundelin J,  Scarborough RM;
```

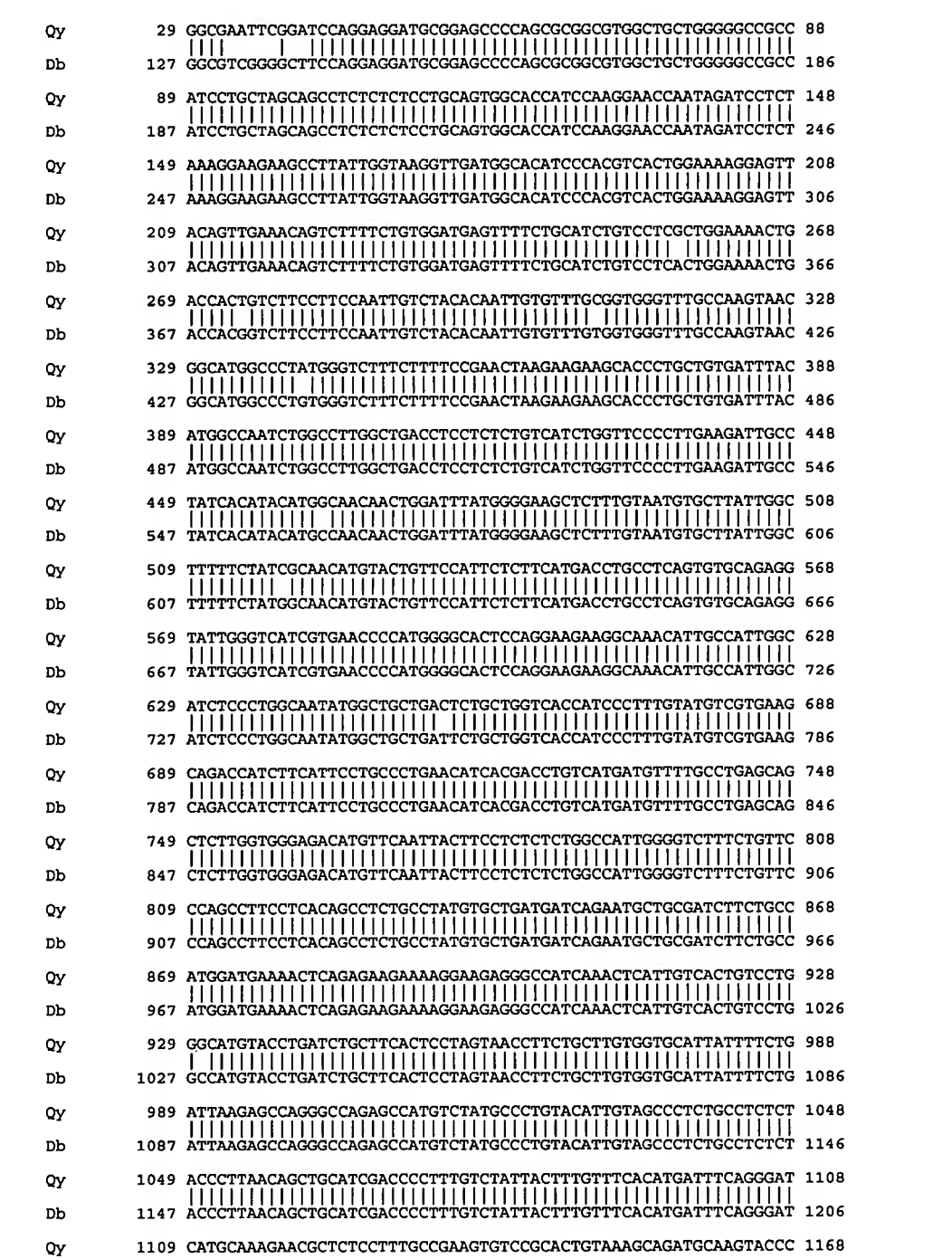
```
XX
DR   WPI; 1996-362813/36.
DR   P-PSDB; AAW01955.
XX
PT   Vector for expression C140 cell surface receptor in host cell - useful to
PT   identify C140 agonist and antagonists, which are antihypertensives and
PT   elevators of blood pressure, respectively.
XX
PS   Example 5; Fig 11A-B; 60pp; English.
XX
CC   AAT32039 encodes the human C140 receptor (C140R). The sequence may be
CC   engineered so as to allow the recombinant expression of C140R in a
CC   suitable host cell, i.e. by removing the native expression-control
CC   sequence and replacing them with control sequences operable in the host.
CC   Such a recombinant receptor can be expressed on the surface of oocytes,
CC   this provides a good assay system for identifying agonists/antagonists of
CC   C140R. The C140 receptor is a G-protein linked receptor and a member of
CC   the "seven-pass" transmembrane receptor superfamily (peptide chain of the
CC   receptor passes through the cell membrane seven times, producing seven
CC   transmembrane regions within the receptor molecule). The C140 receptor is
CC   involved in controlling blood pressure. C140 antagonists (see AAW01942-
CC   W01951) are useful to inhibit signalling from this receptor, resulting in
CC   an increase in blood pressure and are therefore useful in pharmaceuticals
CC   for the treatment of hypotension (low blood pressure). Conversely
CC   agonists (see AAW01914- AAW01941) of C140 are useful in pharmaceuticals
CC   for the treatment of hypertension (high blood pressure)
XX
SQ   Sequence 1414 BP; 335 A; 361 C; 309 G; 409 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 1414;  DB 2;  Length 1414;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1414;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CAAAGAATTGTAATACGACTCACTATAGGGCGAATTCGGATCCAGGAGGATGCGGAGCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAAGAATTGTAATACGACTCACTATAGGGCGAATTCGGATCCAGGAGGATGCGGAGCCC 60

Qy         61 CAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGG 120

Qy        121 CACCATCCAAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACCATCCAAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGG 180

Qy        181 CACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTT 240

Qy        241 TTCTGCATCTGTCCTCGCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCTGCATCTGTCCTCGCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAAT 300

Qy        301 TGTGTTTGCGGTGGGTTTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGTGTTTGCGGTGGGTTTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAAC 360

Qy        361 TAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTC 420

Qy        421 TGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGG 480

Qy        481 GGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCT 540

Qy        541 CTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTC 600

Qy        601 CAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCT 660

Qy        661 GGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCAC 720

Qy        721 GACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCT 780

Qy        781 CTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCT 840

Qy        841 GATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAAGGAAGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAAGGAAGAG 900

Qy        901 GGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACTCCTAGTAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACTCCTAGTAA 960

Qy        961 CCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGC 1020

Qy       1021 CCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTA 1080

Qy       1081 TTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGT 1140

Qy       1141 CCGCACTGTAAAGCAGATGCAAGTACCCCTCACCTCAAAGAAACACTCCAGGAAATCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCGCACTGTAAAGCAGATGCAAGTACCCCTCACCTCAAAGAAACACTCCAGGAAATCCAG 1200

Qy       1201 CTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTTTTCCAGGTCCTCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTTTTCCAGGTCCTCAG 1260

Qy       1261 ATGGGAATTGCACAGTAGGATGTGGAACCTGTTTAATGTTATGAGGACGTGTCTGTTATT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATGGGAATTGCACAGTAGGATGTGGAACCTGTTTAATGTTATGAGGACGTGTCTGTTATT 1320

Qy       1321 TCCGGATCCAGATCTTATTAAAGCAGAACTTGTTTATTGCAGCTTATAATGGTTACAAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TCCGGATCCAGATCTTATTAAAGCAGAACTTGTTTATTGCAGCTTATAATGGTTACAAAT 1380

Qy       1381 AAAGCAATAGCATCACAAATTTCACAAATAAAGC 1414
              |||||||||||||||||||||||||||||||||
Db       1381 AAAGCAATAGCATCACAAATTTCACAAATAAAGC 1414


RESULT 3
ADO28600
ID   ADO28600 standard; cDNA; 2876 BP.
XX
AC   ADO28600;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Human PAR2 encoding cDNA SEQ ID NO:29.
XX
KW   high-grade dysplasia; HGD; oesophageal adenocarcinoma;
KW   neo-plastic transformation; cancer; cytostatic; gene therapy; human;
KW   PAR2; chromosome 5; gene; ss.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   CDS             158. .1348
FT                   /*tag= a
FT                   /product= "PAR2"
FT                   /transl_except= (pos:287. .292,aa:Ser)
FT                   /transl_except= (pos:461. .466,aa:Thr)
FT                   /transl_except= (pos:635. .640,aa:Ser)
FT                   /transl_except= (pos:809. .814,aa:Ala)
FT                   /transl_except= (pos:983. .988,aa:Ser)
FT                   /transl_except= (pos:1157. .1162,aa:Asn)
FT                   /transl_except= (pos:1331. .1336,aa:Lys)
XX
PN   WO2004044178-A2.
XX
PD   27-MAY-2004.
XX
PF   13-NOV-2003; 2003WO-US036260.
XX
PR   13-NOV-2002; 2002US-0425813P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Smith V;
XX
DR   WPI; 2004-420319/39.
DR   P-PSDB; ADO28601.
XX
PT   Detecting of high-grade dysplasia in cells of a mammalian tissue sample
PT   comprises establishing the level of expression in the test tissue sample
PT   of the genes.
XX
PS   Claim 1; SEQ ID NO 29; 256pp; English.
XX
CC   The present invention describes a method for detecting high-grade
CC   dysplasia (HGD) in cells of a mammalian tissue sample. Also described:
CC   (1) identifying an oesophageal tissue susceptible to oesophageal
CC   adenocarcinoma; (2) determining the predisposition of a mammalian tissue
CC   to a neo-plastic transformation by detecting HGD in cells of the tissue;
CC   and (3) detecting cancer in a patient. The method can be used in
CC   detecting HGD and cancer in cells of a mammalian tissue sample. The
CC   methods and compositions of the present invention can be used in treating
CC   and preventing HGD and cancer, and in gene therapy. The present sequence
CC   encodes human PAR2, which is used in the exemplification of the present
CC   invention. The human PAR2 gene is located on chromosome 5.
XX
SQ   Sequence 2876 BP; 772 A; 632 C; 629 G; 843 T; 0 U; 0 Other;

  Query Match           90.5%;  Score 1279.4;  DB 12;  Length 2876;
  Best Local Similarity  97.7%;  Pred. No. 0;
  Matches 1298;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy         29 GGCGAATTCGGATCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 88
              ||||    | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GGCGTCGGGGCTTCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 193

Qy         89 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 253

Qy        149 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 313

Qy        209 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCGCTGGAAAACTG 268
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        314 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTG 373

Qy        269 ACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGTTTGCCAAGTAAC 328
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        374 ACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAAC 433

Qy        329 GGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 388
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        434 GGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 493

Qy        389 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 553

Qy        449 TATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 TATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 613

Qy        509 TTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 568
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 TTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 673

Qy        569 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        674 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 733

Qy        629 ATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 688
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        734 ATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 793

Qy        689 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        794 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 853

Qy        749 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 913

Qy        809 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        914 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 973

Qy        869 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        974 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 1033

Qy        929 GGCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 988
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1034 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 1093

Qy        989 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1094 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1153

Qy       1049 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1154 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1213

Qy       1109 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1214 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCC 1273

Qy       1169 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1333

Qy       1229 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1334 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1393

Qy       1289 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCGGATCCAGATCTTATTAAAGCAGAA 1348
              |||||||||||||||||||||||||||||||||||||||||||||||| ||| | |  | | |  || |
Db       1394 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCTAATCAAAAAGGTCTCACCACATAC 1453

Qy       1349 CTTGTTTAT 1357
              | |||  ||
Db       1454 CATGTGGAT 1462


RESULT 4
AAC60319
```

```
ID   AAC60319 standard; DNA; 1451 BP.
XX
AC   AAC60319;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Human PAR-2 DNA.
XX
KW   PAR-2; protease activated receptor-2; ECL-2; inflammatory disease;
KW   asthma; chronic obstructive pulmonary; arthritis; inflammatory bowel;
KW   psoriasis; eczema; multiple sclerosis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200063371-A1.
XX
PD   26-OCT-2000.
XX
PF   17-APR-2000; 2000WO-GB001455.
XX
PR   15-APR-1999;   99GB-00008513.
XX
PA   (UYSO-) UNIV SOUTHAMPTON.
XX
PI   Walls AF,  Palmer K,  Compton SJ,  Cairns JA,  Gough AC;
XX
DR   WPI; 2000-679599/66.
XX
PT   Protease activated receptor 2 variants useful for treating inflammatory
PT   diseases such as asthma, arthritis and psoriasis, and as hypertensives,
PT   has reduced sensitivity to trypsin.
XX
PS   Disclosure; Page 54-55; 59pp; English.
XX
CC   The present invention relates to a variant protease activated receptor 2
CC   (PAR-2). The invention is useful for identifying an individual having a
CC   polymorphism in the ECL-2 region of one or both PAR-2 gene alleles. The
CC   invention may be used to develop treatments for inflammatory diseases
CC   such as asthma, chronic obstructive pulmonary diseases, arthritis,
CC   inflammatory bowel diseases, psoriasis and eczema, multiple sclerosis and
CC   to raise blood pressure
XX
SQ   Sequence 1451 BP; 310 A; 389 C; 346 G; 406 T; 0 U; 0 Other;

  Query Match             90.0%;  Score 1273;  DB 3;  Length 1451;
  Best Local Similarity   97.7%;  Pred. No. 0;
  Matches 1291;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy         29 GGCGAATTCGGATCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 88
              ||||     | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GGCGTCGGGGCTTCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 186

Qy         89 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 246

Qy        149 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 306

Qy        209 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCGCTGGAAAACTG 268
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        307 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTG 366

Qy        269 ACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGTTTGCCAAGTAAC 328
              ||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        367 ACCACGGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAAC 426

Qy        329 GGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 388
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 486

Qy        389 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 448
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 546

Qy        449 TATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 508
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        547 TATCACATACATGCCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 606

Qy        509 TTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 568
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 TTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 666

Qy        569 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 726

Qy        629 ATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 688
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        727 ATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 786

Qy        689 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 846

Qy        749 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 906

Qy        809 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 966

Qy        869 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 1026

Qy        929 GGCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 988
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 1086

Qy        989 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1146

Qy       1049 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1206

Qy       1109 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1207 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCC 1266

Qy       1169 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1326

Qy       1229 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1386

Qy       1289 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCGGATCCAGATCTTATTAAAGCAGAA 1348
              |||||||||||||||||||||||||||||||||||||| ||| | |  | | |  || |
Db       1387 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCTAATCAAAAAGGTCTCACCACATAC 1446

Qy       1349 C 1349
              |
Db       1447 C 1447


RESULT 5
ABZ42755
ID   ABZ42755 standard; DNA; 1451 BP.
XX
AC   ABZ42755;
```

XX
DT 04-MAR-2003 (first entry)
XX
DE Human proteinase-activated receptor 2 nucleotide SEQ ID NO:299.
XX
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US050107.
XX
PR 19-DEC-2000; 2000US-0257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burmer GC, Roush CL, Brown JP;
XX
DR WPI; 2003-046718/04.
DR P-PSDB; ABP81907.
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.
XX
PS Disclosure; Fig 1; 523pp; English.
XX
CC The present invention describes antigenic peptides (I) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or
CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX
SQ Sequence 1451 BP; 310 A; 389 C; 346 G; 406 T; 0 U; 0 Other;

Query Match 90.0%; Score 1273; DB 8; Length 1451;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 30; Indels 0; Gaps 0;



```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1207 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCC 1266

Qy       1169 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1326

Qy       1229 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1386

Qy       1289 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCGGATCCAGATCTTATTAAAGCAGAA 1348
              ||||||||||||||||||||||||||||||||||||   ||| | |     | | |   || |
Db       1387 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCTAATCAAAAAGGTCTCACCACATAC 1446

Qy       1349 C 1349
              |
Db       1447 C 1447
```

RESULT 6
ADK52593

```
ID   ADK52593 standard; DNA; 1451 BP.
XX
AC   ADK52593;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Hematological disorder associated Gene ID 340.
XX
KW   cytostatic; antianemic; antisickling; virucide; hemostatic; nephrotropic;
KW   cytostatic; thrombolytic; antiparasitic; gene therapy;
KW   hematologic disorder; cancer; Sickle Cell Anemia;
KW   Infectious Mononucleosis; Leukemia; Polycythemia Vera; Lymphoma;
KW   Retinoblastoma; Hemophilia; Thrombosis; Herpes; Thalassemia;
KW   transfusion reaction; Erythroblastosis; mechanical trauma;
KW   micro-angiopathic hemolytic anemia; parasite infection; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             148. .1341
FT                   /*tag=  a
XX
PN   WO2003065871-A2.
XX
PD   14-AUG-2003.
XX
PF   28-JAN-2003; 2003WO-US002484.
XX
PR   04-FEB-2002; 2002US-0354333P.
PR   28-FEB-2002; 2002US-0360258P.
PR   15-MAR-2002; 2002US-0364476P.
PR   26-APR-2002; 2002US-0375626P.
PR   06-JUN-2002; 2002US-0386494P.
PR   24-JUN-2002; 2002US-0390965P.
PR   28-JUN-2002; 2002US-0392480P.
PR   03-JUL-2002; 2002US-0394128P.
PR   31-JUL-2002; 2002US-0399783P.
PR   13-AUG-2002; 2002US-0403221P.
PR   30-AUG-2002; 2002US-0407045P.
PR   25-NOV-2002; 2002US-0429048P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Carroll JM,  Healy A,  Weich NS,  Kelly LM;
XX
DR   WPI; 2003-731464/69.
DR   P-PSDB; ADK52594.
XX
PT   Identifying a compound capable of treating a hematologic disorder (e.g.
PT   anemia or leukemia) comprises assaying the ability of the compound to
PT   modulate the expression or activity of e.g. 131,148, 199 or 12303
PT   polypeptide or nucleic acid.
XX
PS   Disclosure; SEQ ID NO 51; 232pp; English.
XX
CC   The invention relates to a method of identifying a compound capable of
CC   treating a hematologic disorder comprises assaying the ability of the
CC   compound to modulate 131,148, 199, 12303, 13906, 15513, 17822, 302, 5677,
CC   194, 14393, 28059, 7366, 12212, 1981, 261, 12416, 270, 1410, 137, 1871,
CC   13051, 1847, 1849, 15402, 340, 10217, 837, 1761, 8990 or 13249 nucleic
CC   acid expression or polypeptide activity, thus, identifying a compound
CC   capable of treating a hematologic disorder. The methods are useful in
CC   diagnosing, preventing and treating hematological disorders, such as
CC   cancer, Sickle Cell Anemia, Infectious Mononucleosis, Leukemia,
CC   Polycythemia Vera, Lymphoma, Retinoblastoma, Hemophilia, disorders
CC   associated with an increased risk of Thrombosis, Herpes, Thalassemia,
CC   antibody-mediated disorders such as transfusion reactions and
CC   Erythroblastosis, mechanical trauma to red blood cells such as micro-
CC   angiopathic hemolytic anemias, infections by parasites or chemical
CC   injuries. The methods may also be used for identifying compounds that
CC   modulate hematological disorders. This sequence corresponds to one of the
CC   genes modulated the compound.
XX
SQ   Sequence 1451 BP; 310 A; 389 C; 346 G; 406 T; 0 U; 0 Other;

  Query Match             90.0%;  Score 1273;  DB 10;  Length 1451;
  Best Local Similarity   97.7%;  Pred. No. 0;
  Matches 1291;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy         29 GGCGAATTCGGATCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 88
              ||||     |  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GGCGTCGGGGCTTCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 186

Qy         89 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 246

Qy        149 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 306

Qy        209 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCGCTGGAAAACTG 268
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        307 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTG 366

Qy        269 ACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGTTTGCCAAGTAAC 328
              ||||| ||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        367 ACCACGGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAAC 426

Qy        329 GGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 388
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 486

Qy        389 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 546

Qy        449 TATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 508
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        547 TATCACATACATGCCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 606

Qy        509 TTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 568
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 TTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 666

Qy        569 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 726

Qy        629 ATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 688
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        727 ATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 786

Qy        689 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 748
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 846

Qy        749 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 808
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 906

Qy        809 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 868
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 966

Qy        869 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 928
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 1026

Qy        929 GGCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 988
                 | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 1086

Qy        989 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1048
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1146

Qy       1049 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1108
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1206

Qy       1109 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1207 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCC 1266

Qy       1169 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1228
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1326

Qy       1229 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1288
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1386

Qy       1289 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCGGATCCAGATCTTATTAAAGCAGAA 1348
                 |||||||||||||||||||||||||||||||||  ||| | |   | | |   || |
Db       1387 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCTAATCAAAAAGGTCTCACCACATAC 1446

Qy       1349 C 1349
                 |
Db       1447 C 1447
```

RESULT 7
ADN39780
ID ADN39780 standard; cDNA; 1451 BP.
XX
AC ADN39780;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:C152.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerary; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI; 2003-468649/44.
DR P-PSDB; ADN39997.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO C152; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularistaion syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 1451 BP; 310 A; 389 C; 346 G; 406 T; 0 U; 0 Other;

  Query Match   90.0%; Score 1273; DB 11; Length 1451;
  Best Local Similarity   97.7%; Pred. No. 0;
  Matches 1291; Conservative   0; Mismatches   30; Indels   0; Gaps   0;

```
Qy         29 GGCGAATTCGGATCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 88
              ||||    | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GGCGTCGGGGCTTCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 186

Qy         89 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 246

Qy        149 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 306

Qy        209 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCGCTGGAAAACTG 268
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        307 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTG 366
```

```
Qy       269 ACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGTTTGCCAAGTAAC 328
Db       367 ACCACGGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAAC 426

Qy       329 GGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 388
Db       427 GGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 486

Qy       389 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 448
Db       487 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 546

Qy       449 TATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 508
Db       547 TATCACATACATGCCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 606

Qy       509 TTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 568
Db       607 TTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 666

Qy       569 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 628
Db       667 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 726

Qy       629 ATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 688
Db       727 ATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 786

Qy       689 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 748
Db       787 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 846

Qy       749 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 808
Db       847 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 906

Qy       809 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 868
Db       907 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 966

Qy       869 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 928
Db       967 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 1026

Qy       929 GGCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 988
Db      1027 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 1086

Qy       989 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1048
Db      1087 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1146

Qy      1049 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1108
Db      1147 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1206

Qy      1109 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168
Db      1207 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCC 1266

Qy      1169 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1228
Db      1267 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1326

Qy      1229 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1288
Db      1327 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1386

Qy      1289 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCGGATCCAGATCTTATTAAAGCAGAA 1348
Db      1387 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCTAATCAAAAAGGTCTCACCACATAC 1446
```

```
Qy       1349 C 1349
              |
Db       1447 C 1447
```

RESULT 8
ADR46617
ID ADR46617 standard; DNA; 1451 BP.
XX
AC ADR46617;
XX
DT 18-NOV-2004 (first entry)
XX
DE Cancer-associated protein coding sequence, SEQ ID 30.
XX
KW Cytostatic; Gene Therapy; cancer; human; gene; ds.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   148. .1341
FT   /*tag= a
FT   /product= "Cancer-associated protein, SEQ ID 88"
XX
PN WO2004073657-A2.
XX
PD 02-SEP-2004.
XX
PF 19-FEB-2004; 2004WO-US005455.
XX
PR 19-FEB-2003; 2003US-0448784P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Gish KC, Wilson KE, Zlotnik A;
XX
DR WPI; 2004-652787/63.
DR P-PSDB; ADR46675.
XX
PT Detecting a pathological cell in a patient for diagnosing or treating
PT cancer by detecting in a biological sample from the patient genes whose
PT expression are up-regulated or down-regulated in specific cancers.
XX
PS Claim 1; SEQ ID NO 30; 375pp; English.
XX
CC The present invention relates to a method for detecting cancer in a
CC patient. The method comprises detecting in a biological sample from the
CC patient a nucleotide or protein sequence comprising a sequence that is at
CC least 80% identical to a nucleotide sequence (ADR46588-ADR46645) or
CC protein sequence (ADR46646-ADR46703). The method is useful for detecting
CC cancer for preparing a composition for diagnosing or treating cancer.
XX
SQ Sequence 1451 BP; 310 A; 389 C; 346 G; 406 T; 0 U; 0 Other;

  Query Match   90.0%; Score 1273; DB 13; Length 1451;
  Best Local Similarity 97.7%; Pred. No. 0;
  Matches 1291; Conservative   0; Mismatches   30; Indels   0; Gaps   0;

```
Qy         29 GGCGAATTCGGATCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 88
Db        127 GGCGTCGGGGCTTCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 186

Qy         89 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 148
Db        187 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 246

Qy        149 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 208
Db        247 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 306

Qy        209 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCGCTGGAAAACTG 268
Db        307 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTG 366
```

```
Qy        269 ACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGTTTGCCAAGTAAC 328
              ||||| |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        367 ACCACGGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAAC 426

Qy        329 GGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 388
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 486

Qy        389 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 546

Qy        449 TATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 508
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        547 TATCACATACATGCCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 606

Qy        509 TTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 568
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 TTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 666

Qy        569 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 726

Qy        629 ATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 688
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        727 ATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 786

Qy        689 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 846

Qy        749 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 906

Qy        809 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 966

Qy        869 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 1026

Qy        929 GGCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 988
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 1086

Qy        989 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1146

Qy       1049 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1206

Qy       1109 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1207 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCC 1266

Qy       1169 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1326

Qy       1229 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1386

Qy       1289 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCGGATCCAGATCTTATTAAAGCAGAA 1348
              |||||||||||||||||||||||||||||||||  ||| | |   | | |   || |
Db       1387 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCTAATCAAAAAGGTCTCACCACATAC 1446
```

```
Qy       1349 C 1349
              |
Db       1447 C 1447
```

RESULT 9
ADS74019
ID ADS74019 standard; cDNA; 1451 BP.
XX
AC ADS74019;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human G-protein coupled proteinase activated receptor 2 polynucleotide.
XX
KW Human; proteinase activated receptor 2; PAR2; G-protein coupled receptor;
KW receptor; cardiant; neuroprotective; nephrotropic; respiratory-gen. ;
KW gastrointestinal-gen.; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   148. .1341
FT   /*tag= a
FT   /product= "Human PAR2"
XX
PN WO2004080373-A2.
XX
PD 23-SEP-2004.
XX
PF 26-FEB-2004; 2004WO-EP001896.
XX
PR 11-MAR-2003; 2003EP-00004980.
XX
PA (FARB) BAYER HEALTHCARE AG.
XX
PI Golz S, Brueggemeier U, Summer H;
XX
DR WPI; 2004-677358/66.
DR P-PSDB; ADS74020.
XX
PT Screening for therapeutic agents for treating e.g., cardiovascular
PT diseases by contacting a test compound with a proteinase activated
PT receptor 2 (PAR2) polypeptide or polynucleotide and detecting binding of
PT the test compound.
XX
PS Disclosure; SEQ ID NO 1; 121pp; English.
XX
CC The present sequence is that of a polynucleotide encoding G-protein
CC coupled proteinase activated receptor 2 (PAR2). PAR2 is an
CC antiinflammatory receptor in the colon and may also play a role in the
CC airway, regulating sodium ion absorption and anion secretion. The
CC invention relates to novel disease associations of PAR2 polypeptides and
CC polynucleotides. It also relates to novel methods of screening for
CC therapeutic agents for the treatment of cardiovascular disorders,
CC gastrointestinal and liver diseases, neurological disorders, urological
CC disorders, haematological diseases and respiratory diseases in a mammal.
CC Suitable therapeutic agents include a small molecule, an RNA molecule, an
CC antisense oligonucleotide, a polypeptide, an antibody or a ribozyme. The
CC invention also provides pharmaceutical compositions for the treatment of
CC diseases and disorders associated with PAR2 comprising a PAR2
CC polypeptide, PAR2 polynucleotide or a regulator or modulator of PAR2
CC activity. Methods of diagnosing these diseases and disorders involve
CC determining the amount of PAR2 polynucleotide in a sample.
XX
SQ Sequence 1451 BP; 310 A; 389 C; 346 G; 406 T; 0 U; 0 Other;

```
  Query Match             90.0%;  Score 1273;  DB 13;  Length 1451;
  Best Local Similarity   97.7%;  Pred. No. 0;
  Matches 1291;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy         29 GGCGAATTCGGATCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 88
              ||||     | ||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       127 GGCGTCGGGGCTTCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 186

Qy        89 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       187 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 246

Qy       149 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 208
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       247 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTT 306

Qy       209 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCGCTGGAAAACTG 268
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       307 ACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTG 366

Qy       269 ACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGTTTGCCAAGTAAC 328
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       367 ACCACGGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAAC 426

Qy       329 GGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 388
             ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       427 GGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTAC 486

Qy       389 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 448
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       487 ATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCC 546

Qy       449 TATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 508
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       547 TATCACATACATGCCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGC 606

Qy       509 TTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 568
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       607 TTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGG 666

Qy       569 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 628
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       667 TATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGC 726

Qy       629 ATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 688
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       727 ATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 786

Qy       689 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 748
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       787 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAG 846

Qy       749 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 808
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       847 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 906

Qy       809 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 868
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       907 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 966

Qy       869 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 928
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       967 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 1026

Qy       929 GGCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 988
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1027 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTG 1086

Qy       989 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1048
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1087 ATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCT 1146

Qy      1049 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1108
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1147 ACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGAT 1206

Qy      1109 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db      1207 CATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCC 1266

Qy      1169 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1228
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1267 CTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTT 1326

Qy      1229 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1288
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1327 AAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAAC 1386

Qy      1289 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCGGATCCAGATCTTATTAAAGCAGAA 1348
             |||||||||||||||||||||||||||||||||||||||||||  ||| | |   | | |   || |
Db      1387 CTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCTAATCAAAAAGGTCTCACCACATAC 1446

Qy      1349 C 1349
             |
Db      1447 C 1447
```

RESULT 10
ABZ35045
ID ABZ35045 standard; cDNA; 2848 BP.
XX
AC ABZ35045;
XX
DT 05-FEB-2003 (first entry)
XX
DE Human gene expression profile polynucleotide SEQ ID NO 157.
XX
KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
KW gene expression; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200274979-A2.
XX
PD 26-SEP-2002.
XX
PF 20-MAR-2002; 2002WO-US008456.
XX
PR 20-MAR-2001; 2001US-0276947P.
XX
PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI Wan J, Wang Y;
XX
DR WPI; 2002-740862/80.
XX
PT New gene expression profile generated from primary, endothelial,
PT epithelial, and muscle cell types, useful for identifying disease
PT pathologies involving alterations of gene expression, e.g. cancer.
XX
PS Claim 3; Page 381-382; 850pp; English.
XX
CC The invention relates to a gene expression profile comprising one or more
CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
CC is a coronary artery endothelium, umbilical artery or vein endothelium,
CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
CC endothelium, myometrium microvascular endothelium, keratinocyte
CC epithelium, bronchial epithelium, mammary epithelium, prostate
CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
CC small airway epithelium, renal epithelium, umbilical artery smooth
CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
CC osteoblasts or prostate stromal cell. The gene expression profile is used
CC for determining the level of RNA expression for a sample, determining the
CC phenotype of a cell and distinguishing cell types. The gene or a protein
CC expression profile is useful in identifying disease pathologies involving
CC alterations of gene expression. The assessment of expression profiles may
CC provide meaningful information with respect to tumour type and stage,

```
CC   treatment methods, and prognosis. The gene or protein expression profile
CC   may also be used for creating microarrays. The microarray is useful for
CC   genetic and physical mapping of genomes, DNA sequencing, genetic or
CC   medical diagnosis, genotyping of organisms, confirming cell or tissue
CC   identifications and in identifying promising antibiotics, antiviral or
CC   antifungal agents
XX
SQ   Sequence 2848 BP; 670 A; 562 C; 552 G; 771 T; 0 U; 293 Other;

  Query Match             89.4%;  Score 1263.6;  DB 6;  Length 2848;
  Best Local Similarity   97.4%;  Pred. No. 0;
  Matches 1295;  Conservative    0;  Mismatches   34;  Indels    1;  Gaps    1;

Qy         29 GGCGAATTCGGATCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 88
              ||||    |  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 GGCGTCGGGGCTTCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCC 185

Qy         89 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 ATCCTGCTAGCAGCCTCTCTCTCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCT 245

Qy        149 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACT-GGAAAAGGAGT 207
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        246 AAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCACGTCACTGGGAAAAGGAGT 305

Qy        208 TACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCGCTGGAAAACT 267
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        306 TACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACT 365

Qy        268 GACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGTTTGCCAAGTAA 327
              |||||||||||||||||||||||||||||||||||||| |||||| ||  |||||||||
Db        366 GACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGGTTTGCAAGTAA 425

Qy        328 CGGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTA 387
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTA 485

Qy        388 CATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGC 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGC 545

Qy        448 CTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGG 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGG 605

Qy        508 CTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAG 567
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 CTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAG 665

Qy        568 GTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGG 725

Qy        628 CATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCTTTGTATGTCGTGAA 687
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        726 CATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAA 785

Qy        688 GCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCA 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 GCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCA 845

Qy        748 GCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTT 807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        846 GCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTT 905

Qy        808 CCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGC 867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        906 CCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGC 965

Qy        868 CATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCT 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        966 CATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCT 1025

Qy        928 GGGCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCT 987
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 GGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCT 1085

Qy        988 GATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTC 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 GATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTC 1145

Qy       1048 TACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGA 1107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 TACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGA 1205

Qy       1108 TCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACC 1167
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1206 TCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATC 1265

Qy       1168 CCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGT 1227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 CCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGT 1325

Qy       1228 TAAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAA 1287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 TAAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATTGCACAGTAGGATGTGGAA 1385

Qy       1288 CCTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCGGATCCAGATCTTATTAAAGCAGA 1347
              |||||||||||||||||||||||||||||||||||  ||| | |   | | |   || |
Db       1386 CCTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCTAATCAAAAAGGTCTCACCACATA 1445

Qy       1348 ACTTGTTTAT 1357
                | |||  ||
Db       1446 CCATGTGGAT 1455


RESULT 11
AAD44437
ID   AAD44437 standard; DNA; 8624 BP.
XX
AC   AAD44437;
XX
DT   13-DEC-2002  (first entry)
XX
DE   Human coagulation factor II (thrombin) receptor like 1 (F2RL1) gene.
XX
KW   Human; haplotype; coagulation factor II receptor like 1; F2RL1; asthma;
KW   polymorphism; chronic pulmonary disease; inflammatory disorder;
KW   gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(517, G)
FT                   /*tag=  a
FT                   /note= "Polymorphic site; PS1"
FT   variation       replace(553, C)
FT                   /*tag=  b
FT                   /note= "Polymorphic site; PS2"
FT   variation       replace(768, T)
FT                   /*tag=  c
FT                   /note= "Polymorphic site; PS3"
FT   variation       replace(850, T)
FT                   /*tag=  d
FT                   /note= "Polymorphic site; PS4"
FT   variation       replace(852, G)
FT                   /*tag=  e
FT                   /note= "Polymorphic site; PS5"
FT   CDS             1001. .7382
FT                   /*tag=  f
FT                   /product= "Human F2RL1 protein"
FT   exon            1001. .1082
FT                   /*tag=  g
FT                   /number= 1
FT   intron          1083. .6270
FT                   /*tag=  h
```

```
FT   exon            6271. .7382
FT                   /*tag= i
FT                   /number= 2
FT   variation       replace(6277, G)
FT                   /*tag= j
FT                   /note= "Polymorphic site; PS6"
FT   variation       replace(6809, T)
FT                   /*tag= k
FT                   /note= "Polymorphic site; PS7"
FT   variation       replace(7460, C)
FT                   /*tag= l
FT                   /note= "Polymorphic site; PS8"
XX
PN   WO200255534-A2.
XX
PD   18-JUL-2002.
XX
PF   13-NOV-2001; 2001WO-US046475.
XX
PR   10-NOV-2000; 2000US-0247516P.
XX
PA   (GENA-) GENAISSANCE PHARM INC.
XX
PI   Bieglecki KM,  Sanchis A,  Shah N;
XX
DR   WPI; 2002-566728/60.
DR   P-PSDB; AAE26678.
XX
PT   New genetic variants having polymorphisms in the coagulation factor II
PT   (thrombin) receptor like 1 (F2RL1) gene, useful for studying the function
PT   of F2RL1 and treating disorders associated with abnormal expression or
PT   function of F2RL1.
XX
PS   Claim 18; Fig 1; 65pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising genes and
CC   haplotypes of the coagulation factor II (thrombin) receptor like 1
CC   (F2RL1) gene. Polymorphic variants of the F2RL1 gene are useful in
CC   studying the expression and biological function of F2RL1, and in
CC   identifying drugs targetting F2RL1 protein for treating disorders
CC   associated with abnormal expression or function of F2RL1, e.g. asthma,
CC   chronic pulmonary disease, and inflammatory disorders. Polynucleotides
CC   comprising a polymorphic gene variant or fragment may be used for
CC   therapeutic purposes, where a patient could benefit from expression or
CC   increased expression of a particular F2RL1 protein isoform, or an
CC   expression vector encoding the isoform may be administered to the
CC   patient. Haplotype information is useful in improving the efficiency and
CC   output of several steps in drug discovery and development process,
CC   including target validation, identifying lead compounds, and early phase
CC   clinical trials. Information on polymorphisms may be applied in studying
CC   biological functions of F2RL1 as well as in identifying drugs targetting
CC   this protein for the treatment of disorders related to its abnormal
CC   expression or function. The invention is useful in gene therapy. The
CC   present sequence is human F2RL1 gene
XX
SQ   Sequence 8624 BP; 2126 A; 2010 C; 1988 G; 2392 T; 0 U; 108 Other;

  Query Match             84.0%;  Score 1188.4;  DB 6;  Length 8624;
  Best Local Similarity   97.8%;  Pred. No. 0;
  Matches 1201;  Conservative    3;  Mismatches   24;  Indels    0;  Gaps    0;

Qy        130 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 189
              ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6269 AGGAACCARTAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 6328

Qy        190 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6329 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 6388

Qy        250 TGTCCTCGCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGC 309
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6389 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 6448

Qy        310 GGTGGGTTTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAA 369
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       6449 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 6508

Qy        370 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6509 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 6568

Qy        430 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6569 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 6628

Qy        490 TTGTAATGTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGAC 549
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       6629 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 6688

Qy        550 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6689 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 6748

Qy        610 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCAT 669
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       6749 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 6808

Qy        670 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 729
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6809 YCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 6868

Qy        730 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6869 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 6928

Qy        790 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6929 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 6988

Qy        850 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6989 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 7048

Qy        910 ACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 969
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       7049 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 7108

Qy        970 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7109 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 7168

Qy       1030 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCCTTTGTCTATTACTTTGT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7169 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCCTTTGTCTATTACTTTGT 7228

Qy       1090 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7229 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 7288

Qy       1150 AAAGCAGATGCAAGTACCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1209
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       7289 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 7348

Qy       1210 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATT 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7349 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATT 7408

Qy       1270 GCACAGTAGGATGTGGAACCTGTTTAATGTTATGAGGACGTGTCTGTTATTTCCGGATCC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||:||  |||||||
Db       7409 GCACAGTAGGATGTGGAACCTGTTTAATGTTATGAGGACGTGTCTGTTATTYCCTAATCA 7468

Qy       1330 AGATCTTATTAAAGCAGAACTTGTTTAT 1357
              | |  | | |  || | | |||  ||
Db       7469 AAAAGGTCTCACCACATACCATGTGGAT 7496
```

```
RESULT 12
AAD44438
ID   AAD44438 standard; DNA; 1194 BP.
XX
AC   AAD44438;
XX
DT   13-DEC-2002  (first entry)
XX
DE   Human coagulation factor II (thrombin) receptor like 1 (F2RL1) DNA.
XX
KW   Human; haplotype; coagulation factor II receptor like 1; F2RL1; asthma;
KW   polymorphism; chronic pulmonary disease; inflammatory disorder;
KW   gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1194
FT                   /*tag=  a
FT                   /product= "Human F2RL1 protein"
FT   variation       replace(89, G)
FT                   /*tag=  b
FT                   /note= "Polymorphic site"
FT   variation       replace(621, T)
FT                   /*tag=  c
FT                   /note= "Polymorphic site"
XX
PN   WO200255534-A2.
XX
PD   18-JUL-2002.
XX
PF   13-NOV-2001; 2001WO-US046475.
XX
PR   10-NOV-2000; 2000US-0247516P.
XX
PA   (GENA-) GENAISSANCE PHARM INC.
XX
PI   Bieglecki KM,  Sanchis A,  Shah N;
XX
DR   WPI; 2002-566728/60.
DR   P-PSDB; AAE26678.
XX
PT   New genetic variants having polymorphisms in the coagulation factor II
PT   (thrombin) receptor like 1 (F2RL1) gene, useful for studying the function
PT   of F2RL1 and treating disorders associated with abnormal expression or
PT   function of F2RL1.
XX
PS   Claim 23; Fig 2; 65pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising genes and
CC   haplotypes of the coagulation factor II (thrombin) receptor like 1
CC   (F2RL1) gene. Polymorphic variants of the F2RL1 gene are useful in
CC   studying the expression and biological function of F2RL1, and in
CC   identifying drugs targetting F2RL1 protein for treating disorders
CC   associated with abnormal expression or function of F2RL1, e.g. asthma,
CC   chronic pulmonary disease, and inflammatory disorders. Polynucleotides
CC   comprising a polymorphic gene variant or fragment may be used for
CC   therapeutic purposes, where a patient could benefit from expression or
CC   increased expression of a particular F2RL1 protein isoform, or an
CC   expression vector encoding the isoform may be administered to the
CC   patient. Haplotype information is useful in improving the efficiency and
CC   output of several steps in drug discovery and development process,
CC   including target validation, identifying lead compounds, and early phase
CC   clinical trials. Information on polymorphisms may be applied in studying
CC   biological functions of F2RL1 as well as in identifying drugs targetting
CC   this protein for the treatment of disorders related to its abnormal
CC   expression or function. The invention is useful in gene therapy. The
CC   present sequence is human F2RL1 DNA
XX
SQ   Sequence 1194 BP; 264 A; 321 C; 261 G; 348 T; 0 U; 0 Other;

  Query Match             83.6%;  Score 1182.8;  DB 6;  Length 1194;
  Best Local Similarity   99.4%;  Pred. No. 0;

  Matches 1187;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         50 ATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCAGCCTCTCTC 109
Db          1 ATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCAGCCTCTCTC 60

Qy        110 TCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGT 169
Db         61 TCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGT 120

Qy        170 AAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCT 229
Db        121 AAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCT 180

Qy        230 GTGGATGAGTTTTCTGCATCTGTCCTCGCTGGAAAACTGACCACTGTCTTCCTTCCAATT 289
Db        181 GTGGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATT 240

Qy        290 GTCTACACAATTGTGTTTGCGGTGGGTTTGCCAAGTAACGGCATGGCCCTATGGGTCTTT 349
Db        241 GTCTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTT 300

Qy        350 CTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCT 409
Db        301 CTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCT 360

Qy        410 GACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAAC 469
Db        361 GACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAAC 420

Qy        470 TGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATCGCAACATGTAC 529
Db        421 TGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTAC 480

Qy        530 TGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCC 589
Db        481 TGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCC 540

Qy        590 ATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 649
Db        541 ATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 600

Qy        650 CTGACTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 709
Db        601 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 660

Qy        710 CTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTC 769
Db        661 CTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTC 720

Qy        770 AATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCT 829
Db        721 AATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCT 780

Qy        830 GCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAG 889
Db        781 GCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAG 840

Qy        890 AAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTC 949
Db        841 AAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTC 900

Qy        950 ACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGC 1009
Db        901 ACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGC 960

Qy       1010 CATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGAC 1069
Db        961 CATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGAC 1020

Qy       1070 CCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTT 1129
Db       1021 CCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTT 1080
```

```
Qy       1130 TGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCCCTCACCTCAAAGAAACACTCC 1189
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1081 TGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCC 1140

Qy       1190 AGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGA 1243
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGA 1194


RESULT 13
AAZ50773
ID   AAZ50773 standard; DNA; 1289 BP.
XX
AC   AAZ50773;
XX
DT   31-MAY-2000  (first entry)
XX
DE   Human protease activated receptor-2 DNA.
XX
KW   Human; PAR-2; protease activated receptor; antisense molecule;
KW   PAR antibody; cytostatic; therapeutic; metastatic tumour cell;
KW   placental implantation; invasive cell; ds.
XX
OS   Homo sapiens.
XX
PN   WO200008150-A1.
XX
PD   17-FEB-2000.
XX
PF   05-FEB-1999;   99WO-IL000079.
XX
PR   07-AUG-1998;   98IL-00125698.
XX
PA   (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX
PI   Bar-Shavit R;
XX
DR   WPI; 2000-205706/18.
XX
PT   Treating metastatic tumor cells useful for treating disorders involving
PT   placenta implantation in a female comprises administration of an
PT   antisense molecule complementary to an RNA sequence of a protease
PT   activated receptor protein.
XX
PS   Example 3; Fig 9; 46pp; English.
XX
CC   The patent discloses a method to treat metastatic tumour cells using an
CC   antisense molecule comprising a polynucleotide complementary to an RNA
CC   sequence of a protease activated receptor (PAR) protein, or an antibody
CC   capable of binding to a PAR protein. The antisense molecules and
CC   antibodies of PAR protein are also used to treat disorders associated
CC   with implantation of placenta. The present sequence is a human PAR-2 DNA
CC   used for producing antisense molecules for treating invasive cells
XX
SQ   Sequence 1289 BP; 300 A; 320 C; 270 G; 399 T; 0 U; 0 Other;

  Query Match           83.6%;  Score 1182.8;  DB 3;  Length 1289;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1187;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        130 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 154

Qy        190 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 214

Qy        250 TGTCCTCGCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGC 309
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 274

Qy        310 GGTGGGTTTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAA 369
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        275 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 334

Qy        370 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 394

Qy        430 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 454

Qy        490 TTGTAATGTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGAC 549
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        455 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 514

Qy        550 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        515 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 574

Qy        610 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCAT 669
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        575 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 634

Qy        670 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        635 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 694

Qy        730 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        695 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 754

Qy        790 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        755 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 814

Qy        850 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        815 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 874

Qy        910 ACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 969
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        875 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 934

Qy        970 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        935 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 994

Qy       1030 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        995 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1054

Qy       1090 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1055 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1114

Qy       1150 AAAGCAGATGCAAGTACCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1209
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1115 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1174

Qy       1210 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATT 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1175 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTTTTCCAGGTCCTCAGATGGGAATT 1234

Qy       1270 GCACAGTAGGATGTGGAACCTGTTTAATGTTATGAGGACGTGTCTGTTATTTCC 1323
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1235 GCACAGTAGGATGTGGAACCTGTTTAATGTTATGAGGACGTGTCTGTTATTTCC 1288


RESULT 14
ADO29874
ID   ADO29874 standard; cDNA; 1194 BP.
XX
```

AC ADO29874;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human GPCR F2RL1 polynucleotide, SEQ ID NO:976.
XX
KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antimanic;
KW cytostatic; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW dermatological; antiulcer; antithyroid; antiallergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004040000-A2.
XX
PD 13-MAY-2004.
XX
PF 09-SEP-2003; 2003WO-US028226.
XX
PR 09-SEP-2002; 2002US-0409303P.
PR 09-APR-2003; 2003US-0461329P.
XX
PA (PRIM-) PRIMAL INC.
XX
PI Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
PI Madisen L, Mcilwain KL, Pavlova MN, Vassilatis D, Zeng H;
XX
DR WPI; 2004-390329/36.
DR P-PSDB; ADO29311.
XX
PT Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
PS Claim 151; SEQ ID NO 976; 542pp; English.
XX
CC The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the trasngenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridise to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC nucleic acid of the invention. Note: The full sequence data for this
CC patent did not form part of the printed specification; those sequences
CC not shown were obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1194 BP; 264 A; 322 C; 261 G; 347 T; 0 U; 0 Other;

Query Match 83.4%; Score 1179.6; DB 12; Length 1194;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy      50 ATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCAGCCTCTCTC 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCAGCCTCTCTC 60

Qy     110 TCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGT 169
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TCCTGCAGTGGCACCATCCAAGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGT 120

Qy     170 AAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCT 229
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AAGGTTGATGGCACATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCT 180

Qy     230 GTGGATGAGTTTTCTGCATCTGTCCTCGCTGGAAAACTGACCACTGTCTTCCTTCCAATT 289
           |||||||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db     181 GTGGATGAGTTTTCTGCATCTGTCCTCACTGGAAAACTGACCACGGTCTTCCTTCCAATT 240

Qy     290 GTCTACACAATTGTGTTTGCGGTGGGTTTGCCAAGTAACGGCATGGCCCTATGGGTCTTT 349
           |||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||||
Db     241 GTCTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTT 300

Qy     350 CTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCT 409
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTTTTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCT 360

Qy     410 GACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAAC 469
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     361 GACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGCCAACAAC 420

Qy     470 TGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATCGCAACATGTAC 529
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     421 TGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTAC 480

Qy     530 TGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCC 589
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCC 540

Qy     590 ATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 649
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ATGGGGCACTCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 600

Qy     650 CTGACTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 709
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 660

Qy     710 CTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTC 769
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CTGAACATCACGACCTGTCATGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTC 720

Qy     770 AATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCT 829
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 AATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCT 780

Qy     830 GCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAG 889
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAG 840

Qy     890 AAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTC 949
           |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db     841 AAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTC 900
```

```
Qy       950 ACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGC 1009
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 ACTCCTAGTAACCTTCTGCTTGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGC 960

Qy      1010 CATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGAC 1069
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 CATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGAC 1020

Qy      1070 CCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTT 1129
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 CCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTT 1080

Qy      1130 TGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCCCTCACCTCAAAGAAACACTCC 1189
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      1081 TGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCC 1140

Qy      1190 AGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGA 1243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 AGGAAATCCAGCTCTTACTCTTCAAGTTCAACCACTGTTAAGACCTCCTATTGA 1194


RESULT 15
AAQ84558
ID   AAQ84558 standard; DNA; 1255 BP.
XX
AC   AAQ84558;
XX
DT   25-MAR-2003  (revised)
DT   22-AUG-1995  (first entry)
XX
DE   Human C140 receptor genomic DNA.
XX
KW   G-protein-coupled receptor; G-protein; C140 receptor; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             56. .1197
FT                   /*tag= a
XX
PN   WO9503318-A1.
XX
PD   02-FEB-1995.
XX
PF   26-JUL-1994;   94WO-US008536.
XX
PR   26-JUL-1993;   93US-00097938.
XX
PA   (CORT-) COR THERAPEUTICS.
XX
PI   Scarborough RM,  Sundelin J;
XX
DR   WPI; 1995-075182/10.
DR   P-PSDB; AAR66921.
XX
PT   New DNA encoding recombinant C140 receptor - and novel agonists and
PT   antagonists and specific antibodies with therapeutic and diagnostic
PT   applications.
XX
PS   Disclosure; Fig 2; 57pp; English.
XX
CC   The availability of genomic DNA encoding the mouse protease C140 receptor
CC   (see Q84557) permitted the retrieval of the corresp. human gene. A human
CC   genomic library cloned in the vector EMBL3 was screened using the entire
CC   coding region of the murine clone as a probe. The recovered human gene
CC   including the DNA sequence and the deduced AA sequence are shown in
CC   Q84558 & R66921. Subsequent experiments indicated that the human C140
CC   gene is located in the same region of the long arm of chromosome number 5
CC   (5q12-5q13) as has been reported for the human thrombin receptor gene.
CC   *(Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 1255 BP; 294 A; 320 C; 260 G; 381 T; 0 U; 0 Other;


  Query Match           78.2%;  Score 1105.8;  DB 2;  Length 1255;
  Best Local Similarity  99.4%;  Pred. No. 0;
  Matches 1110;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       130 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 189
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       139 AGGAACCAATAGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACATCCCA 198

Qy       190 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 249
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       199 CGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATC 258

Qy       250 TGTCCTCGCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGC 309
            ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       259 TGTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGT 318

Qy       310 GGTGGGTTTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAA 369
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       319 GGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAA 378

Qy       370 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 429
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       379 GCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTG 438

Qy       430 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 489
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       439 GTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCT 498

Qy       490 TTGTAATGTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGAC 549
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       499 TTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGAC 558

Qy       550 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 609
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       559 CTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGGAAGAA 618

Qy       610 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCAT 669
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       619 GGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCAT 678

Qy       670 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 729
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       679 CCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCA 738

Qy       730 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 789
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       739 TGATGTTTTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGC 798

Qy       790 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 849
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       799 CATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAG 858

Qy       850 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 909
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       859 AATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAA 918

Qy       910 ACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 969
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       919 ACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCT 978

Qy       970 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1029
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       979 TGTGGTGCATTATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACAT 1038

Qy      1030 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1089
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1039 TGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGT 1098

Qy      1090 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1149
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1099 TTCACATGATTTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGT 1158

Qy      1150 AAAGCAGATGCAAGTACCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1209
```

```
                ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1159 AAAGCAGATGCAAGTATCCCTCACCTCAAAGAAACACTCCAGGAAATCCAGCTCTTACTC 1218

Qy       1210 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1246
                ||||||||||||||||||||||||||||||||||||
Db       1219 TTCAAGTTCAACCACTGTTAAGACCTCCTATTGAGTT 1255


Search completed: March 21, 2005, 18:18:56
Job time : 830.407 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 20:59:39 ; Search time 68.9132 Seconds
(without alignments)
2228.076 Million cell updates/sec

Title: US-10-643-627-63
Perfect score: 2031
Sequence: 1 MRSPSAAWLLGAAILLAASL.....KHSRKSSSYSSSSTTVKTSY 397

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2023	99.6	397	2	AAR66923	Aar66923 Human C14
2	2023	99.6	397	2	AAW01955	Aaw01955 Human C14
3	2003	98.6	397	3	AAB35641	Aab35641 Human PAR
4	2003	98.6	397	5	AAE26678	Aae26678 Human coa
5	2003	98.6	397	6	ABG73508	Abg73508 Human par
6	2003	98.6	397	7	ADE62812	Ade62812 Human Pro
7	2003	98.6	397	8	ADO29311	Ado29311 Human GPC
8	2003	98.6	397	8	ADS74020	Ads74020 Human G-p
9	1998	98.4	397	8	ADL61221	Adl61221 Human coa
10	1997	98.3	397	6	ABP81907	Abp81907 Human pro
11	1997	98.3	397	7	ADK52594	Adk52594 Hematolog
12	1997	98.3	397	7	ADN39997	Adn39997 Cancer/an
13	1997	98.3	397	8	ADR46675	Adr46675 Cancer-as
14	1965.5	96.8	394	2	AAW51408	Aaw51408 Human pro
15	1884.5	92.8	389	8	ADO28601	Ado28601 Human PAR
16	1880	92.6	398	2	AAR66921	Aar66921 Human C14
17	1880	92.6	398	2	AAW01953	Aaw01953 Human C14
18	1793	88.3	355	8	ADI28653	Adi28653 Human mod
19	1720	84.7	341	8	ADI28654	Adi28654 Human mod
20	1713	84.3	399	2	AAR66922	Aar66922 Murine C1
21	1713	84.3	399	2	AAW01954	Aaw01954 Murine C1
22	1713	84.3	399	7	ABR63562	Abr63562 Delayed h
23	1713	84.3	399	8	ADO29312	Ado29312 Mouse GPC
24	1707	84.0	397	7	ADE62810	Ade62810 Rat Prote
25	1644.5	81.0	395	2	AAR66920	Aar66920 Murine C1
26	1639.5	80.7	395	2	AAW01952	Aaw01952 Murine C1
27	1625	80.0	320	8	ADI28655	Adi28655 Human mod
28	611	30.1	420	6	ABG73510	Abg73510 X. laevis
29	609	30.0	430	8	ADO29310	Ado29310 Mouse GPC
30	608	29.9	425	2	AAR27240	Aar27240 Human thr
31	608	29.9	425	2	AAR60698	Aar60698 Fragment
32	608	29.9	425	2	AAW51407	Aaw51407 Human pro
33	608	29.9	425	2	AAY49570	Aay49570 Human thr
34	608	29.9	425	5	AAE17032	Aae17032 Human thr
35	608	29.9	425	5	AAG80697	Aag80697 Human thr
36	608	29.9	425	6	ABG73511	Abg73511 Human thr
37	608	29.9	425	6	ABR47449	Abr47449 Breast ca
38	608	29.9	425	6	ABP81919	Abp81919 Human thr
39	608	29.9	425	7	ADG89876	Adg89876 Human coa
40	608	29.9	425	8	ADL14208	Adl14208 Novel hum
41	608	29.9	425	8	ADN04016	Adn04016 Antipsori
42	608	29.9	425	8	ADO29309	Ado29309 Human GPC
43	608	29.9	425	8	ADQ18985	Adq18985 Human sof
44	608	29.9	425	8	ADR45608	Adr45608 Human G p
45	608	29.9	425	8	ADS84489	Ads84489 Human pro

ALIGNMENTS

RESULT 1
AAR66923
ID AAR66923 standard; protein; 397 AA.
XX
AC AAR66923;
XX
DT 25-MAR-2003 (revised)
DT 22-AUG-1995 (first entry)
XX
DE Human C140 receptor encoded by cDNA.
XX
KW G-protein-coupled receptor; G-protein; C140 receptor.
XX
OS Homo sapiens.
XX
PN WO9503318-A1.
XX
PD 02-FEB-1995.
XX
PF 26-JUL-1994; 94WO-US008536.
XX
PR 26-JUL-1993; 93US-00097938.
XX
PA (CORT-) COR THERAPEUTICS.
XX
PI Scarborough RM, Sundelin J;
XX
DR WPI; 1995-075182/10.
DR N-PSDB; AAQ84560.
XX
PT New DNA encoding recombinant C140 receptor - and novel agonists and
PT antagonists and specific antibodies with therapeutic and diagnostic
PT applications.
XX
PS Example; Fig 11; 57pp; English.
XX
CC A human intestinal tumour cDNA library was subjected to PCR using primers
CC designed from the genomic clone (see AAQ84558) and the amplified fragment
CC was cloned in pSG5 and sequenced. There are four AA differences between
CC the cDNA encoded sequence and that encoded by the genomic DNA. The
CC genomic DNA sequence and deduced AA sequence are given in AAQ84560 &
CC AAR66923. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 397 AA;

Query Match 99.6%; Score 2023; DB 2; Length 397;
Best Local Similarity 99.7%; Pred. No. 1.1e-210;
Matches 396; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYRNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||||||||||||||||||||||||||||||
Db        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 2
AAW01955
ID AAW01955 standard; protein; 397 AA.
XX
AC AAW01955;
XX
DT 02-APR-1997 (first entry)
XX
DE Human C140 receptor.
XX
KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .27
FT /note= "the signal peptide differs from that encoded by a
FT genomic DNA sequence for this receptor (see AAW01953),
FT the signal sequence given here is believed to be the
FT correct sequence"
FT Protein 28. .397
FT /note= "mature protein"
XX
PN WO9623225-A1.
XX
PD 01-AUG-1996.
XX
PF 25-JAN-1996; 96WO-US001179.
XX
PR 25-JAN-1995; 95US-00390301.
XX
PA (CORT-) COR THERAPEUTICS INC.
XX
PI Sundelin J, Scarborough RM;
XX
DR WPI; 1996-362813/36.
DR N-PSDB; AAT32039.
XX
PT Vector for expression C140 cell surface receptor in host cell - useful to
PT identify C140 agonist and antagonists, which are antihypertensives and
PT elevators of blood pressure, respectively.
XX
PS Example 5; Fig 11A-B; 60pp; English.
XX
CC AAW01955 represents the human C140 receptor (C140R). DNA encoding C140R
CC may be engineered so as to allow the recombinant expression of C140R in a
CC suitable host cell, i.e. by removing the native expression-control
CC sequences and replacing them with control sequences operable in the host.
CC Such a recombinant receptor can be expressed on the surface of oocytes,
CC this provides a good assay system for identifying agonists/antagonists of
CC C140R. The C140 receptor is a G-protein linked receptor and a member of
CC the "seven-pass" transmembrane receptor superfamily (peptide chain of the
CC receptor passes through the cell membrane seven times, producing seven
CC transmembrane regions within the receptor molecule). The C140 receptor is
CC involved in controlling blood pressure. C140 antagonists (see AAW01942-
CC W01951) are useful to inhibit signalling from this receptor, resulting in
CC an increase in blood pressure and are therefore useful in pharmaceuticals
CC for the treatment of hypotension (low blood pressure). Conversely
CC agonists (see AAW01914-W01941) of C140 are useful in pharmaceuticals for
CC the treatment of hypertension (high blood pressure)
XX
SQ Sequence 397 AA;

Query Match 99.6%; Score 2023; DB 2; Length 397;
Best Local Similarity 99.7%; Pred. No. 1.1e-210;
Matches 396; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYRNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||||||||||||||||||||||||||||||
Db        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 3
AAB35641
ID AAB35641 standard; protein; 397 AA.
XX
AC AAB35641;
XX
DT 19-FEB-2001 (first entry)
XX
DE Human PAR-2 protein.
XX
KW PAR-2; protease activated receptor-2; ECL-2; inflammatory disease;
KW asthma; chronic obstructive pulmonary; arthritis; inflammatory bowel;
KW psoriasis; eczema; multiple sclerosis.
XX
OS Homo sapiens.
XX
PN WO200063371-A1.
XX
PD 26-OCT-2000.

XX
PF 17-APR-2000; 2000WO-GB001455.
XX
PR 15-APR-1999; 99GB-00008513.
XX
PA (UYSO-) UNIV SOUTHAMPTON.
XX
PI Walls AF, Palmer K, Compton SJ, Cairns JA, Gough AC;
XX
DR WPI; 2000-679599/66.
XX
PT Protease activated receptor 2 variants useful for treating inflammatory
PT diseases such as asthma, arthritis and psoriasis, and as hypertensives,
PT has reduced sensitivity to trypsin.
XX
PS Claim 2; Page 55; 59pp; English.
XX
CC The present invention relates to a variant protease activated receptor 2
CC (PAR-2). The invention is useful for identifying an individual having a
CC polymorphism in the ECL-2 region of one or both PAR-2 gene alleles. The
CC invention may be used to develop treatments for inflammatory diseases
CC such as asthma, chronic obstructive pulmonary diseases, arthritis,
CC inflammatory bowel diseases, psoriasis and eczema, multiple sclerosis and
CC to raise blood pressure
XX
SQ Sequence 397 AA;

Query Match 98.6%; Score 2003; DB 3; Length 397;
Best Local Similarity 98.7%; Pred. No. 1.7e-208;
Matches 392; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
Db 1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy 61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
Db 61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy 121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
Db 121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy 181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
Db 181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy 241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
Db 241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 300

Qy 301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
Db 301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy 361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
Db 361 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 397

RESULT 4
AAE26678
ID AAE26678 standard; protein; 397 AA.
XX
AC AAE26678;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human coagulation factor II (thrombin) receptor like 1 (F2RL1) protein.
XX
KW Human; haplotype; coagulation factor II receptor like 1; F2RL1; asthma;
KW polymorphism; chronic pulmonary disease; inflammatory disorder;
KW gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 30
FT /note= "This amino acid changes to Ser due to single
FT nucleotide polymorphism"
XX
PN WO200255534-A2.
XX
PD 18-JUL-2002.
XX
PF 13-NOV-2001; 2001WO-US046475.
XX
PR 10-NOV-2000; 2000US-0247516P.
XX
PA (GENA-) GENAISSANCE PHARM INC.
XX
PI Bieglecki KM, Sanchis A, Shah N;
XX
DR WPI; 2002-566728/60.
DR N-PSDB; AAD44437, AAD44438.
XX
PT New genetic variants having polymorphisms in the coagulation factor II
PT (thrombin) receptor like 1 (F2RL1) gene, useful for studying the function
PT of F2RL1 and treating disorders associated with abnormal expression or
PT function of F2RL1.
XX
PS Claim 27; Fig 3; 65pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising genes and
CC haplotypes of the coagulation factor II (thrombin) receptor like 1
CC (F2RL1) gene. Polymorphic variants of the F2RL1 gene are useful in
CC studying the expression and biological function of F2RL1, and in
CC identifying drugs targetting F2RL1 protein for treating disorders
CC associated with abnormal expression or function of F2RL1, e.g. asthma,
CC chronic pulmonary disease, and inflammatory disorders. Polynucleotides
CC comprising a polymorphic gene variant or fragment may be used for
CC therapeutic purposes, where a patient could benefit from expression or
CC increased expression of a particular F2RL1 protein isoform, or an
CC expression vector encoding the isoform may be administered to the
CC patient. Haplotype information is useful in improving the efficiency and
CC output of several steps in drug discovery and development process,
CC including target validation, identifying lead compounds, and early phase
CC clinical trials. Information on polymorphisms may be applied in studying
CC biological functions of F2RL1 as well as in identifying drugs targetting
CC this protein for the treatment of disorders related to its abnormal
CC expression or function. The invention is useful in gene therapy. The
CC present sequence is human F2RL1 protein
XX
SQ Sequence 397 AA;

Query Match 98.6%; Score 2003; DB 5; Length 397;
Best Local Similarity 98.7%; Pred. No. 1.7e-208;
Matches 392; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
Db 1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy 61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
Db 61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy 121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
Db 121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy 181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
Db 181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy 241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
                ||||||||||||| ||||||||||||||||||||||||
Db        361 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 397


RESULT 5
ABG73508
ID   ABG73508 standard; protein; 397 AA.
XX
AC   ABG73508;
XX
DT   14-FEB-2003  (first entry)
XX
DE   Human par2 protein SEQ ID 39.
XX
KW   G-protein coupled receptor; HGPRBMY1; HGPRBMY2; immunosuppressive;
KW   cardiant; neuroprotective; antiinflammatory; cytostatic; vulnerary;
KW   vaccine; gene therapy; autoimmune; cardiovascular; neural; reproductive;
KW   haematopoietic; pulmonary; gastrointestinal; proliferation; cell cycle;
KW   birth defect; aberrant phosphorylation; acute phase response; receptor;
KW   signal transduction; hyperimmune activity; inflammatory; hypercongenital;
KW   necrotic lesion; wound; organ transplant rejection.
XX
OS   Homo sapiens.
XX
PN   WO200268591-A2.
XX
PD   06-SEP-2002.
XX
PF   22-FEB-2002; 2002WO-US005281.
XX
PR   23-FEB-2001; 2001US-0270792P.
PR   23-FEB-2001; 2001US-0270793P.
PR   06-JUN-2001; 2001US-0296427P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Feder J,  Ramanathan C,  Nelson T,  Mintier G,  Cacace A,  Barber L;
PI   Kornacker M,  Bol D;
XX
DR   WPI; 2003-058304/05.
XX
PT   New human HGPRBMY1 or HGPRBMY2 polynucleotide and polypeptide, useful
PT   preventing, treating or ameliorating a disorder e.g., wound,
PT   cardiovascular disorder or transplant rejection.
XX
PS   Disclosure; Fig 4; 316pp; English.
XX
CC   This invention describes the novel human G-protein coupled receptors
CC   (GPCR's), HGPRBMY1 or HGPRBMY2 which have immunosuppressive, cardiant,
CC   neuroprotective, antiinflammatory, cytostatic and vulnerary activity and
CC   can be used in vaccines or for gene therapy. Pharmaceutical compositions
CC   comprising HGPRBMY1 or HGPRBMY2 polypeptides or their agonists or
CC   antagonists or modulators, or a HGPRBMY1- or HGPRBMY2-specific antibody
CC   are useful for preventing, treating or ameliorating a medical condition
CC   comprising autoimmune, cardiovascular, neural, reproductive,
CC   haematopoietic, pulmonary, gastrointestinal or proliferating disorder, a
CC   cell cycle or birth defect, a disorder related to aberrant
CC   phosphorylation, acute phase responses or signal transduction or to
CC   hyperimmune activity, an inflammatory or hypercongenital condition, a
CC   necrotic lesion, a wound, organ transplant rejection or a condition
CC   related to organ transplant rejection. This sequence represents a G-
CC   protein coupled receptor associated with the human HGPRBMY proteins
CC   described in the disclosure of the invention
XX
SQ   Sequence 397 AA;

  Query Match             98.6%;  Score 2003;  DB 6;  Length 397;
  Best Local Similarity   98.7%;  Pred. No. 1.7e-208;
  Matches  392;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              ||||||||||||| ||||||||||||||||||||||||
Db        361 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 397


RESULT 6
ADE62812
ID   ADE62812 standard; protein; 397 AA.
XX
AC   ADE62812;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human Protein P55085, SEQ ID NO 8745.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; P55085.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
```

```
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a human protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 397 AA;

  Query Match             98.6%;  Score 2003;  DB 7;  Length 397;
  Best Local Similarity   98.7%;  Pred. No. 1.7e-208;
  Matches  392;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||||| |||||||||||||||||||||||||
Db        361 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 397


RESULT 7
ADO29311
ID   ADO29311 standard; protein; 397 AA.
XX
AC   ADO29311;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human GPCR F2RL1, SEQ ID NO:412.
XX
KW   G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW   transgenic mouse; neurological disorder; adrenal gland disorder;
KW   colon disorder; intestinal disorder; cardiovascular disorder;
KW   muscular disorder; blood disorder; immune disorder; bone disorder;
KW   joint disorder; metabolic disorder; nutritive disorder; cancer;
KW   kidney disorder; liver disorder; lung disorder; breast disorder;
KW   ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW   skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW   thymus disorder; thyroid disorder; antiparkinsonian; antimanic;
KW   cytostatic; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
KW   CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW   virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW   dermatological; antiulcer; antithyroid; antiallergic; anorectic;
KW   immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW   receptor.
XX
OS   Homo sapiens.
XX
PN   WO2004040000-A2.
XX
PD   13-MAY-2004.
XX
PF   09-SEP-2003; 2003WO-US028226.
XX
PR   09-SEP-2002; 2002US-0409303P.
PR   09-APR-2003; 2003US-0461329P.
XX
PA   (PRIM-) PRIMAL INC.
XX
PI   Gaitanaris GA,  Bergmann JE,  Gragerov A,  Hohmann J,  Li F;
PI   Madisen L,  Mcilwain KL,  Pavlova MN,  Vassilatis D,  Zeng H;
XX
DR   WPI; 2004-390329/36.
DR   N-PSDB; ADO29874.
XX
PT   Novel mammalian G protein coupled receptors, useful for identifying
PT   compounds that modulates diagnosing and treating disease condition
PT   associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT   pectoris, Parkinson's disease.
XX
PS   Claim 151; SEQ ID NO 412; 542pp; English.
XX
CC   The invention relates to human and mouse G protein-coupled receptors
CC   (GPCRs) and nucleic acids encoding them. The invention also relates to
CC   sequences at least 90% identical to the GPCR proteins and nucleic acids
CC   of the invention; methods of treating, preventing or diagnosing diseases
CC   associated with GPCRs of the invention; methods of screening for
CC   compounds useful in the treatment of GPCR-related diseases; a transgenic
CC   mouse comprising a GPCR gene of the invention; a mouse comprising a
CC   mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC   from the trasngenic mice; kits comprising several mice, each of which has
CC   a mutation in a different GPCR gene of the invention; and kits comprising
CC   probes which hybridise to GPCR polynucleotides of the invention. The
CC   invention further discloses variants of the GPCR polypeptides and vectors
CC   comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC   be used in the diagnosis, treatment or prevention of a wide variety of
CC   diseases including neurological disorders (e.g., Alzheimer's disease,
CC   depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC   disorders of the adrenal gland; disorders of the colon or intestine
CC   (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC   syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC   myocardial infarction); muscular disorders; blood disorders (e.g.,
CC   anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC   AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC   arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC   obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC   diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC   uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC   thyroid (e.g., cancers). The present sequence represents a GPCR of the
CC   invention. Note: The full sequence data for this patent did not form part
CC   of the printed specification; those sequences not shown were obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
```

XX
SQ Sequence 397 AA;

Query Match 98.6%; Score 2003; DB 8; Length 397;
Best Local Similarity 98.7%; Pred. No. 1.7e-208;
Matches 392; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||||| |||||||||||||||||||||||
Db        361 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 8
ADS74020
ID ADS74020 standard; protein; 397 AA.
XX
AC ADS74020;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human G-protein coupled proteinase activated receptor 2 (PAR2).
XX
KW Human; proteinase activated receptor 2; PAR2; G-protein coupled receptor;
KW receptor; cardiant; neuroprotective; nephrotropic; respiratory-gen.;
KW gastrointestinal-gen.; gene therapy.
XX
OS Homo sapiens.
XX
PN WO2004080373-A2.
XX
PD 23-SEP-2004.
XX
PF 26-FEB-2004; 2004WO-EP001896.
XX
PR 11-MAR-2003; 2003EP-00004980.
XX
PA (FARB) BAYER HEALTHCARE AG.
XX
PI Golz S, Brueggemeier U, Summer H;
XX
DR WPI; 2004-677358/66.
DR N-PSDB; ADS74019.
XX
PT Screening for therapeutic agents for treating e.g., cardiovascular
PT diseases by contacting a test compound with a proteinase activated
PT receptor 2 (PAR2) polypeptide or polynucleotide and detecting binding of
PT the test compound.
XX
PS Disclosure; SEQ ID NO 2; 121pp; English.
XX
CC The present sequence is that of human G-protein coupled proteinase
CC activated receptor 2 (PAR2). PAR2 is an antiinflammatory receptor in the
CC colon and may also play a role in the airway, regulating sodium ion
CC absorption and anion secretion. The invention relates to novel disease
CC associations of PAR2 polypeptides and polynucleotides. It also relates to
CC novel methods of screening for therapeutic agents for the treatment of
CC cardiovascular disorders, gastrointestinal and liver diseases,
CC neurological disorders, urological disorders, haematological diseases and
CC respiratory diseases in a mammal. Suitable therapeutic agents include a
CC small molecule, an RNA molecule, an antisense oligonucleotide, a
CC polypeptide, an antibody or a ribozyme. The invention also provides
CC pharmaceutical compositions for the treatment of diseases and disorders
CC associated with PAR2 comprising a PAR2 polypeptide, PAR2 polynucleotide
CC or a regulator or modulator of PAR2 activity. Methods of diagnosing these
CC diseases and disorders involve determining the amount of PAR2
CC polynucleotide in a sample.
XX
SQ Sequence 397 AA;

Query Match 98.6%; Score 2003; DB 8; Length 397;
Best Local Similarity 98.7%; Pred. No. 1.7e-208;
Matches 392; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||||| |||||||||||||||||||||||
Db        361 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 9
ADL61221
ID ADL61221 standard; protein; 397 AA.
XX
AC ADL61221;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human coagulation factor II (thrombin) receptor-like 1 protein.
XX
KW predictor set; protein tyrosine kinase; cytostatic; antiangiogenic;
KW vasotropic; vulnerary; pharmacogenomic; drug sensitivity; breast cancer;
KW hypervascular disease; angiogenesis; wound healing scar; human;
KW biomarker; coagulation factor II receptor-like 1; thrombin; receptor.
XX
OS Homo sapiens.
XX
PN WO2004020583-A2.
XX
PD 11-MAR-2004.

XX
PF 26-AUG-2003; 2003WO-US026491.
XX
PR 27-AUG-2002; 2002US-0406385P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Huang F, Han X, Reeves KA, Amler L, Fairchild CR, Lee FY;
PI Shaw P;
XX
DR WPI; 2004-239171/22.
DR N-PSDB; ADL61084.
XX
PT New predictor sets with a plurality of polynucleotides and/or
PT polypeptides whose expression pattern predicts cell response to a
PT compound that modulates protein tyrosine kinase activity, useful in
PT treating breast cancer.
XX
PS Claim 9; SEQ ID NO 145; 649pp; English.
XX
CC The invention relates to a novel predictor set comprising a plurality of
CC polynucleotides and/or polypeptides whose expression pattern is
CC predictive of the response of cells to treatment with a compound that
CC modulates protein tyrosine kinase activity or members of the protein
CC tyrosine kinase pathway. The molecules of the invention demonstrate
CC cytostatic, antiangiogenic, vasotropic and vulnerary activities and may
CC be useful in the field of pharmacogenomics, in particular for determining
CC drug sensitivity and in treating breast cancer, hypervascular diseases,
CC angiogenesis and scars in wound healing. The current sequence is that of
CC a human protein tyrosine kinase biomarker protein of the invention.
XX
SQ Sequence 397 AA;

Query Match 98.4%; Score 1998; DB 8; Length 397;
Best Local Similarity 98.5%; Pred. No. 5.9e-208;
Matches 391; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVPS 60
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTSRSSKGRSLIGKVDGTSHVTGKGVTVETVPS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||||| |||||||||||||||||||||||
Db        361 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 10
ABP81907
ID ABP81907 standard; protein; 397 AA.
XX
AC ABP81907;
XX
DT 04-MAR-2003 (first entry)
XX
DE Human proteinase-activated receptor 2 protein SEQ ID NO:300.
XX
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer.
XX
OS Homo sapiens.
XX
PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US050107.
XX
PR 19-DEC-2000; 2000US-0257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burmer GC, Roush CL, Brown JP;
XX
DR WPI; 2003-046718/04.
DR N-PSDB; ABZ42755.
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.
XX
PS Disclosure; Fig 1; 523pp; English.
XX
CC The present invention describes antigenic peptides (I) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or
CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX
SQ Sequence 397 AA;

Query Match 98.3%; Score 1997; DB 6; Length 397;
Best Local Similarity 98.5%; Pred. No. 7.6e-208;
Matches 391; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVPS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHANNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||||| |||||||||||||||||||||||||
Db        361 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 11
ADK52594
ID ADK52594 standard; protein; 397 AA.
XX
AC ADK52594;
XX
DT 06-MAY-2004 (first entry)
XX
DE Hematological disorder associated Gene ID 340 encoded protein.
XX
KW cytostatic; antianemic; antisickling; virucide; hemostatic; nephrotropic;
KW cytostatic; thrombolytic; antiparasitic; gene therapy;
KW hematologic disorder; cancer; Sickle Cell Anemia;
KW Infectious Mononucleosis; Leukemia; Polycythemia Vera; Lymphoma;
KW Retinoblastoma; Hemophilia; Thrombosis; Herpes; Thalassemia;
KW transfusion reaction; Erythroblastosis; mechanical trauma;
KW micro-angiopathic hemolytic anemia; parasite infection.
XX
OS Homo sapiens.
XX
PN WO2003065871-A2.
XX
PD 14-AUG-2003.
XX
PF 28-JAN-2003; 2003WO-US002484.
XX
PR 04-FEB-2002; 2002US-0354333P.
PR 28-FEB-2002; 2002US-0360258P.
PR 15-MAR-2002; 2002US-0364476P.
PR 26-APR-2002; 2002US-0375626P.
PR 06-JUN-2002; 2002US-0386494P.
PR 24-JUN-2002; 2002US-0390965P.
PR 28-JUN-2002; 2002US-0392480P.
PR 03-JUL-2002; 2002US-0394128P.
PR 31-JUL-2002; 2002US-0399783P.
PR 13-AUG-2002; 2002US-0403221P.
PR 30-AUG-2002; 2002US-0407045P.
PR 25-NOV-2002; 2002US-0429048P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Carroll JM, Healy A, Weich NS, Kelly LM;
XX
DR WPI; 2003-731464/69.
DR N-PSDB; ADK52593.
XX
PT Identifying a compound capable of treating a hematologic disorder (e.g.
PT anemia or leukemia) comprises assaying the ability of the compound to
PT modulate the expression or activity of e.g. 131,148, 199 or 12303
PT polypeptide or nucleic acid.
XX
PS Disclosure; SEQ ID NO 52; 232pp; English.
XX
CC The invention relates to a method of identifying a compound capable of
CC treating a hematologic disorder comprises assaying the ability of the
CC compound to modulate 131,148, 199, 12303, 13906, 15513, 17822, 302, 5677,
CC 194, 14393, 28059, 7366, 12212, 1981, 261, 12416, 270, 1410, 137, 1871,
CC 13051, 1847, 1849, 15402, 340, 10217, 837, 1761, 8990 or 13249 nucleic
CC acid expression or polypeptide activity, thus, identifying a compound
CC capable of treating a hematologic disorder. The methods are useful in
CC diagnosing, preventing and treating hematological disorders, such as
CC cancer, Sickle Cell Anemia, Infectious Mononucleosis, Leukemia,
CC Polycythemia Vera, Lymphoma, Retinoblastoma, Hemophilia, disorders
CC associated with an increased risk of Thrombosis, Herpes, Thalassemia,
CC antibody-mediated disorders such as transfusion reactions and
CC Erythroblastosis, mechanical trauma to red blood cells such as micro-
CC angiopathic hemolytic anemias, infections by parasites or chemical
CC injuries. The methods may also be used for identifying compounds that
CC modulate hematological disorders. This sequence corresponds to the
CC protein encoded by one of the genes modulated by the compounds.
XX
SQ Sequence 397 AA;

Query Match 98.3%; Score 1997; DB 7; Length 397;
Best Local Similarity 98.5%; Pred. No. 7.6e-208;
Matches 391; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHANNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||||| |||||||||||||||||||||||||
Db        361 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 12
ADN39997
ID ADN39997 standard; protein; 397 AA.
XX
AC ADN39997;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C367.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;

KW retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerary; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI; 2003-468649/44.
DR N-PSDB; ADN39780.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 12; SEQ ID NO C367; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularistaion syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
SQ Sequence 397 AA;

Query Match 98.3%; Score 1997; DB 7; Length 397;
Best Local Similarity 98.5%; Pred. No. 7.6e-208;
Matches 391; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHANNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||||| |||||||||||||||||||||||
Db        361 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 13
ADR46675
ID ADR46675 standard; protein; 397 AA.
XX
AC ADR46675;
XX
DT 18-NOV-2004 (first entry)
XX
DE Cancer-associated protein, SEQ ID 88.
XX
KW Cytostatic; Gene Therapy; cancer; human.
XX
OS Homo sapiens.
XX
PN WO2004073657-A2.
XX
PD 02-SEP-2004.
XX
PF 19-FEB-2004; 2004WO-US005455.
XX
PR 19-FEB-2003; 2003US-0448784P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Gish KC, Wilson KE, Zlotnik A;
XX
DR WPI; 2004-652787/63.
DR N-PSDB; ADR46617.
XX
PT Detecting a pathological cell in a patient for diagnosing or treating
PT cancer by detecting in a biological sample from the patient genes whose
PT expression are up-regulated or down-regulated in specific cancers.
XX
PS Claim 1; SEQ ID NO 88; 375pp; English.
XX
CC The present invention relates to a method for detecting cancer in a
CC patient. The method comprises detecting in a biological sample from the
CC patient a nucleotide or protein sequence comprising a sequence that is at
CC least 80% identical to a nucleotide sequence (ADR46588-ADR46645) or
CC protein sequence (ADR46646-ADR46703). The method is useful for detecting
CC cancer for preparing a composition for diagnosing or treating cancer.
XX
SQ Sequence 397 AA;

Query Match 98.3%; Score 1997; DB 8; Length 397;
Best Local Similarity 98.5%; Pred. No. 7.6e-208;
Matches 391; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHANNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 300

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||||| ||||||||||||||||||||||||
Db        361 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 397
```

RESULT 14
AAW51408
ID AAW51408 standard; protein; 394 AA.
XX
AC AAW51408;
XX
DT 12-OCT-1998 (first entry)
XX
DE Human protease-activated receptor 2 (PAR2).
XX
KW Protease-activated receptor 2; PAR2; PAR3; thrombin receptor; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Cleavage-site 36. .37
FT /note= "thrombin cleavage site"
XX
PN WO9818456-A1.
XX
PD 07-MAY-1998.
XX
PF 29-OCT-1997; 97WO-US019732.
XX
PR 30-OCT-1996; 96US-00742440.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Coughlin SR, Ishihara H, Connolly A;
XX
DR WPI; 1998-271905/24.
XX
PT DNA encoding protease-activated receptor 3 - for detection of specific
PT agonists and antagonists, potentially useful for treating e.g.
PT thrombosis, atherosclerosis, inflammation etc.
XX
PS Example 1; Page 43-44; 74pp; English.
XX
CC This polypeptide comprises human protease-activated receptor 2 (PAR2).
CC The physiological activator of PAR2 remains unknown; it is not activated
CC by thrombin. The invention relates to novel mouse and human PAR3 (see
CC AAW51405-06) that show homology to PAR2 and which are specific receptors
CC for thrombin. They can be used to screen for specific agonists and
CC antagonists of thrombin useful e.g. for treating atherosclerosis,
CC thrombosis and inflammation
XX
SQ Sequence 394 AA;

Query Match 96.8%; Score 1965.5; DB 2; Length 394;
Best Local Similarity 97.7%; Pred. No. 2e-204;
Matches 388; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

```
Qy          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFS 60

Qy         61 VDEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120
              ||||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 VDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALA 120

Qy        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        121 DLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 180

Qy        181 MGHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMF 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 MGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDP- 239

Qy        241 NYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        240 --FLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICF 297

Qy        301 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 TPSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALL 357

Qy        361 CRSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||||| ||||||||||||||||||||||||
Db        358 CRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY 394
```

RESULT 15
ADO28601
ID ADO28601 standard; protein; 389 AA.
XX
AC ADO28601;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human PAR2 protein SEQ ID NO:30.
XX
KW high-grade dysplasia; HGD; oesophageal adenocarcinoma;
KW neo-plastic transformation; cancer; cytostatic; gene therapy; human;
KW PAR2; chromosome 5.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 44
FT /note= "encoded by ACATCC"
FT Misc-difference 101
FT /note= "encoded by CGAACT"
FT Misc-difference 158
FT /note= "encoded by TGTTCC"
FT Misc-difference 215
FT /note= "encoded by CCTGCC"
FT Misc-difference 272
FT /note= "encoded by AACTCA"
FT Misc-difference 329
FT /note= "encoded by CTTAAC"
FT Misc-difference 386
FT /note= "encoded by GTTAAG"
XX
PN WO2004044178-A2.
XX
PD 27-MAY-2004.

XX
PF 13-NOV-2003; 2003WO-US036260.
XX
PR 13-NOV-2002; 2002US-0425813P.
XX
PA (GETH) GENENTECH INC.
XX
PI Smith V;
XX
DR WPI; 2004-420319/39.
DR N-PSDB; ADO28600.
XX
PT Detecting of high-grade dysplasia in cells of a mammalian tissue sample
PT comprises establishing the level of expression in the test tissue sample
PT of the genes.
XX
PS Disclosure; SEQ ID NO 30; 256pp; English.
XX
CC The present invention describes a method for detecting high-grade
CC dysplasia (HGD) in cells of a mammalian tissue sample. Also described:
CC (1) identifying an oesophageal tissue susceptible to oesophageal
CC adenocarcinoma; (2) determining the predisposition of a mammalian tissue
CC to a neo-plastic transformation by detecting HGD in cells of the tissue;
CC and (3) detecting cancer in a patient. The method can be used in
CC detecting HGD and cancer in cells of a mammalian tissue sample. The
CC methods and compositions of the present invention can be used in treating
CC and preventing HGD and cancer, and in gene therapy. The present sequence
CC represents human PAR2, which is used in the exemplification of the
CC present invention. The human PAR2 gene is located on chromosome 5.
XX
SQ Sequence 389 AA;

Query Match 92.8%; Score 1884.5; DB 8; Length 389;
Best Local Similarity 97.0%; Pred. No. 1.3e-195;
Matches 384; Conservative 0; Mismatches 5; Indels 7; Gaps 7;

```
Qy          2 RSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSV 61
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 RSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG-SHVTGKGVTVETVFSV 59

Qy         62 DEFSASVLAGKLTTVFLPIVYTIVFAVGLPSNGMALWVFLFRTKKKHPAVIYMANLALAD 121
              |||||||| |||||||||||||||| |||||||||||||| ||||||||||||||||||||
Db         60 DEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLF-TKKKHPAVIYMANLALAD 118

Qy        122 LLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPM 181
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        119 LLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMY-SILFMTCLSVQRYWVIVNPM 177

Qy        182 GHSRKKANIAIGISLAIWLLTLLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFN 241
              |||||||||||||||||| |||||||||||||||| |||||||||||||||||||||||||
Db        178 GHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFI-ALNITTCHDVLPEQLLVGDMFN 236

Qy        242 YPLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLGMYLICFT 301
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||| |||||||
Db        237 YPLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDE-SEKKRKRAIKLIVTVLAMYLICFT 295

Qy        302 PSNLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLC 361
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        296 PSNLLLVVHYFLIKSQGQSHVYALYIVALCLST-NSCIDPFVYYFVSHDFRDHAKNALLC 354

Qy        362 RSVRTVKQMQVPLTSKKHSRKSSSYSSSSTTVKTSY 397
              |||||||||| |||||||||||||||||| ||||
Db        355 RSVRTVKQMQVSLTSKKHSRKSSSYSSSSTT-KTSY 389
```

Search completed: March 18, 2005, 21:06:44
Job time : 69.9132 secs